



NATIONAL BANK OF KAZAKHSTAN

## Implementation of artificial intelligence in the NBRK





# AI - is the basic technology that will be implemented in all key processes.



## Analytics and forecasting

Financial stability, forecast rounds



*In development*



## Financial stability and security

AML / fraud / Cyber



*In development*



## Operational efficiency

NLP, document management, assistants, search



*In development*



## Customer and service solutions

Bots, services for participants



*during the analysis process*



## Supervision and regulation

RegTech / SupTech, automatic reporting analysis



*during the analysis process*

### Supports

01

Data Fabric

02

GPU

03

The AI pipeline

### Responsible persons



AI Champion



The owner of the process

### The key principle



Effective interaction between technical and business teams is the basis for successful implementation of AI solutions in a corporate environment.



# Analytics and Forecasting: **AN INTELLIGENT FINANCIAL ANALYST**

Creating a digital twin that answers questions and integrates all corporate analytics for proactive decision-making



## The client part

### AI Agent Interface



## Workflow

### Workflow engine



Process orchestration

### Governance

VerifyWise /  
Permite.io

Verifying the rights to use the AI  
platform

### Pre-processing tools



Guardrails  
Protection and control when  
interacting with LLM



## LLM

### Base LLM

GPT oss. 120B

### Embedding models

Nomic-embed-text  
mxbai-embed-large

Creating vector  
representations of data

### Vector Store

Supabase / Chroma

Storing vector  
representations of data



## Data sources / domains

### Data domains

- Balance of payments, EIVK, CTR
- ESSP, FRSP, kred.reg., dep.reg.
- Fin.household. activity of the real sector of the economy, assets, liabilities, real estate registry, KASE

### Data transmission

- Denodo (API)

### Current status:

- The **data source defined** within the **MVP** implementation is **credit register**.
- The **LLM model** will connect to the data source via the Denodo API and perform tokenization.

### Expected result:

- ✓ **Scenario modeling**: answers to questions, scenarios for modeling situations, transparency of interpretation of thoughts and methodological basis
- ✓ **Price stabilization**: recommending necessary measures, simulating the consequences of decisions before they are made.
- ✓ **A single analytics window**: Access to all corporate analytics and sources in one place: data factory, GNI, reporting, analytics.



# Financial stability and security: IDENTIFICATION OF SHADOW TRANSFERS

A model to detect covert transfers used for money laundering or the movement of criminal proceeds through sole proprietors (individual entrepreneurs) and personal accounts, with the aim of strengthening internal controls.

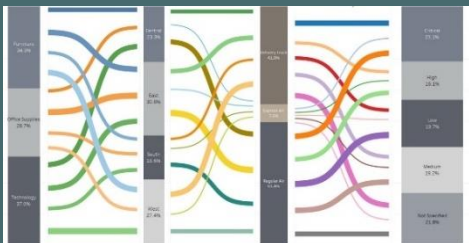


## The client part

### Interaction Interface

Qlik

SANKEY DIAGRAM



## ML и Workflow

### Chains of connections of questionable transactions have been built

- Clustering of anomalous nodes based on DBSCAN or Isolation Forest followed by analysis via PCA

### Workflow engine

Flowise



## Data sources / domains

### Data source

- Report on completed foreign exchange transactions: PR-9 (ESSP)

### Data preparation

- Creating masked data
- Implementing a transactional graph
- Forming node features

### Current status:

The **constructed graph** enables automatic detection of suspicious transactions and relationships. An algorithm for constructing chains of relationships has been developed. To improve accuracy, it is planned to use DBSCAN or Isolation Forest models to identify patterns of suspicious transactions.

### Expected result:

- ✓ A classification algorithm will identify all possible patterns.
- ✓ Based on the analysis, all chains that can be identified as suspicious patterns of shadow transfers are identified.
- ✓ Following the confirmed labeling, a report will be sent to DFMVK business users for analysis based on the relationship diagram.



# Operational Efficiency: **DIGITAL ASSISTANT**

A chatbot for NBRK employees to quickly search for necessary information and generate analytical reports within the framework of existing regulations, rules, instructions, and conducted research.



## The client part

### AI Agent Interface



## Workflow

### Workflow engine



Flowise

Process orchestration

### Governance

VerifyWise /  
Permite.ioVerifying the rights to use the AI  
platform

### Pre-processing tools



Guardrails

Protection and control when  
interacting with LLM

## LLM

### Base LLM

GPT oss. 120B

### Embedding models

Nomic-embed-text  
mx-bai-embed-largeCreating vector  
representations of data

### Vector Store

Supabase / Chroma

Storing vector  
representations of data

## Data sources / domains

### Data domains

- Rules, orders, regulations
- Reports, studies, economic reviews, international standards,

### Current status:

- The HR's open source data (rules, regulations, guidelines, agreements, and laws) have been compiled.
- A MVP chatbot has been developed based on the collected HR open source of data.
- A design prototype has been developed on Outsystems.
- We are awaiting the receipt of the infrastructure for deploying the LLM model.

### Expected result:

- ✓ Reduce information search time by 30%
- ✓ Optimize routine tasks with a digital assistant

**IT dep.**  
PC and phone loan  
Computer peripheral loan

**Security dep.**  
Rules and safety  
precautions

**HR**

**Own structural  
division**  
Regulations  
Rules



# Internal AI Consultant of JSC GKB

A chatbot for employees of JSC GKB for the prompt search of necessary information within the framework of existing regulations, rules, instructions, etc.



## Client side



## Workflow



## LLM



## Data sources/domains

### AI Agent Interface Internal website (React)

#### AI Router

t-tech (T-lite-it-1.0)  
(Ha бaзe Qwen 2.5)

Distributes incoming requests (AI router)

#### Protection and control

Closed perimeter – the model and vector storage operate only within the GKB internal network.

System prompt – strictly prohibits responses outside of the found fragments.

#### Base LLM

t-tech/T-lite-it-1.0  
(Ha бaзe Qwen 2.5)

#### Embedding models

FRIDA (ai-forever)      Creating vector representations of data  
BGE-reranker (BAAI)

#### Vector Store

Milvus      Storing vector representations of data

#### Data domains

- Rules, orders, regulations, and VND
- Reference Books
- Technical Documentation and Knowledge Base (Confluence)

### Current status:

- 200 VND of JSC GKB (rules, regulations, guidelines, agreements, laws) were collected.
- One release based on the collected VND data and a company directory was developed.
- A design prototype was developed in React.
- Infrastructure (GPU) is required to improve the quality and scalability of the LLM model.

### Expected result::

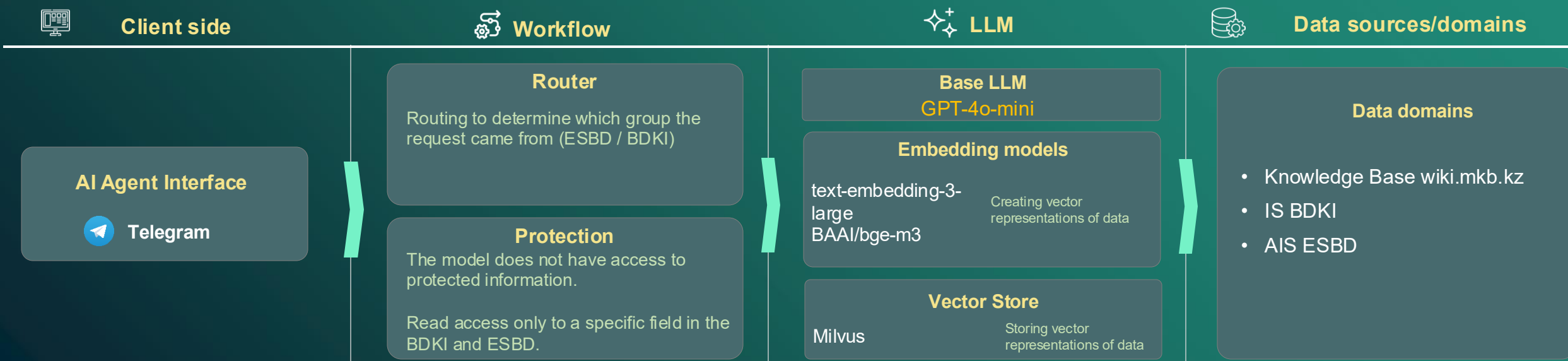
- ✓ Reduce information search time by 30%
- ✓ Optimize routine tasks with a digital assistant





# External AI Consultant of JSC GKB - Arman AI

Creating an AI assistant for suppliers based on the wiki.mkb.kz knowledge base to provide assistance on various issues.



## Current status:

- An automatic check for data uploads to the BDKI has been developed.
- A standalone assistant for checking the status of push notifications for insurance contracts has been developed.

## Expected result:

Unloading and assistance to OKP (service desk) staff



# Analytical initiatives of JSC GKB

Creation of national digital ratings for insurance and credit assessment of citizens



## Client side



## Rating



## Data sources

### Ratings interface for citizens



eGov Mobile

id.mkb.kz

Arman Bot

### Ratings interface for the financial sector API

### Insurance rating

A digital indicator for assessing insurance reliability, improving financial literacy, and encouraging participation in voluntary insurance

MLOps, LightGBM, XGBoost, SHAP/LIME, Ensemble models, Optuna

### Credit Rating + Recommendation Model\*

Creation of a digital profile of a citizen with an individual credit rating and a system of personal recommendations to improve the indicator

MLOps, LightGBM, Ensemble models, PD, Optuna, SHAP/LIME

### Data domains

- IS BDKI
- AIS ESBD
- HD GBD

### Current status:

**Insurance Rating** - Behavioral Maturity Assessments (testing), Insurance Risk Assessment model under development.

**Credit Rating** - Logic and Calculation Indicators Testing

### Expected Insurance Rating Result:

This digital indicator of client reliability in the insurance industry helps foster a culture of voluntary insurance and motivates responsible financial behavior.

### Expected Credit Rating Result:

Allows citizens not only to track their credit history but also to receive steps to improve and restore it.





# A unified platform is the basis for all AI agents of the Anti-Fraud Center



## Interfaces and interactions

**Orchestration of dialogues and actions**  
LangGraph

**Service layer and integrations**  
FastAPI

**Caching**  
Redis



## Processing and orchestration

**Managing all pipelines**  
DolphinScheduler

**Exchange of events between services**  
Kafka

**Stream data processing**  
Flink

**Extracting text and structure from documents**  
Tika, Tesseract и Camelot



## LLM

### Self-hosted LLM

- A universal model for analyzing, understanding, and generating texts
- Tasks: classification, fact extraction, report and response generation, NER, translation, summarization
- Working via vLLM



## Security and observability

**Access control**  
Keycloak/Casdoor

**Politicians and Secrets**  
Politicians - OPA,  
secrets - Vault

**Monitoring and metrics**  
Prometheus + Grafana

**Tracing**  
OpenTelemetry

**LLM Quality Control and Logging**  
Langfuse/RAGAS



## Data sources

**Data Center - Selena Lakehouse**

- A unified platform for storing and processing data, including documents, metadata, and vector representations
- Built-in vector search

**To construct and analyze a fraudulent connection graph**  
ArangoDB

**For full-text search**  
OpenSearch



### Note.

During implementation, the architecture and tools used may be adjusted.



## Analytics and forecasting : **OBSERVER AGENT**

Creating an agent to monitor fraudulent schemes on the internet, the darknet, and social media.

MVP: Analyzing industry reports and identifying fraud trends

### STAGE 1 (MVP)

Analyzing industry reports and identifying fraud trends

### STAGE 2

- 1) Integration of the agent with the Anti-Fraud Center
- 2) Social media analysis for the following purposes:
  - identifying fraud trends;
  - data leaks; DDoS attacks

### STAGE 3

Darknet analysis for the following purposes:

- identifying fraud trends;
- data leaks; DDoS attacks

### **Expected result:**

Fraud Trends: Identifying fraud trends based on data analysis



## Improving Efficiency: AGENT - ANALYST

Creating an agent that will generate analytical reports on request  
MVP: Generating reports and relationship graphs

### STAGE 1 (MVP)

Creation of an analytical report based on data from the Anti-Fraud Center, with recommendations and instructions on further actions

### STAGE 2

- 1) Integrating the agent with the Anti-Fraud Center
- 2) Analyzing the connections between incidents and incident parameters and other incidents, creating connection graphs

### Expected result:

Optimization of the work of criminal prosecution authorities



## Process Optimization: SUPPORT AGENT

Creating an agent that accompanies users of the system

### STAGE 1

Support chat, providing answers to frequently asked questions, searching for answers in instructions, FAQs, and technical documentation

### STAGE 2

Поддержка по работе с интерфейсом Антифрод-центра (поиск, фильтры, заполнение инцидентов и др.)

#### **Expected result:**

Improve service for Antifraud Center participants and provide support for working within the Antifraud Center system. Free up resources.



## Process optimization: AGENT - OPERATOR

Creation of an agent that provides support to entities listed by the Anti-Fraud Center

### STAGE 1

Support chat, providing answers to frequently asked questions, searching for answers in instructions, FAQ

### STAGE 2

Integration of the agent with the Antifraud Center and the Central Information Center

### Expected result:

Improve service for entities included in the Anti-Fraud Center databases and provide support. Free up resources.