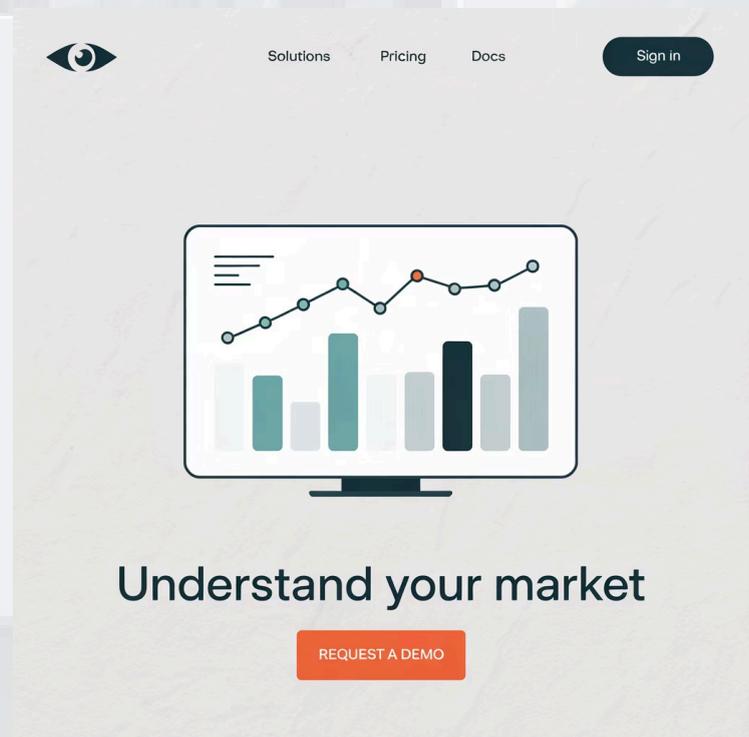


# Comparing Competitors' Finished Products for Pricing Strategy

## Client and Challenge

**Client:** Large full-cycle metallurgical company

**Problem:** Comparing characteristics of finished products with competitor analogs for accurate pricing. Disparate data formats and names made manual analysis extremely time-consuming - the process took days.



01

### ML pattern matching model

Development of an algorithm for automatic matching of product batches by chemical composition, fraction, and quality indicators

30

minutes

Time to prepare a report instead of several days

02

### Data Normalization

Creation of a unified nomenclature reference and standardization of data formats from various suppliers

03

### Interactive Dashboard

Development of a panel displaying price dynamics and automatic recommendations for the commercial department

5-7%

margin growth

On key contracts thanks to accurate pricing

✔ **Key Achievement:** The system ensured a quick reaction to changes in competitor prices, which is critically important in the metallurgical industry

**Technologies:** Python Pandas • scikit-learn • FastAPI • PostgreSQL • Plotly Dash

# LLM + RAG for Document Search

A corporation with thousands of documents in various formats needed a single tool for quick information retrieval.

## Problem

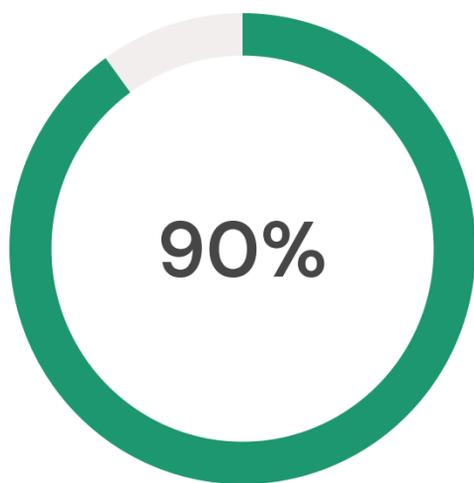
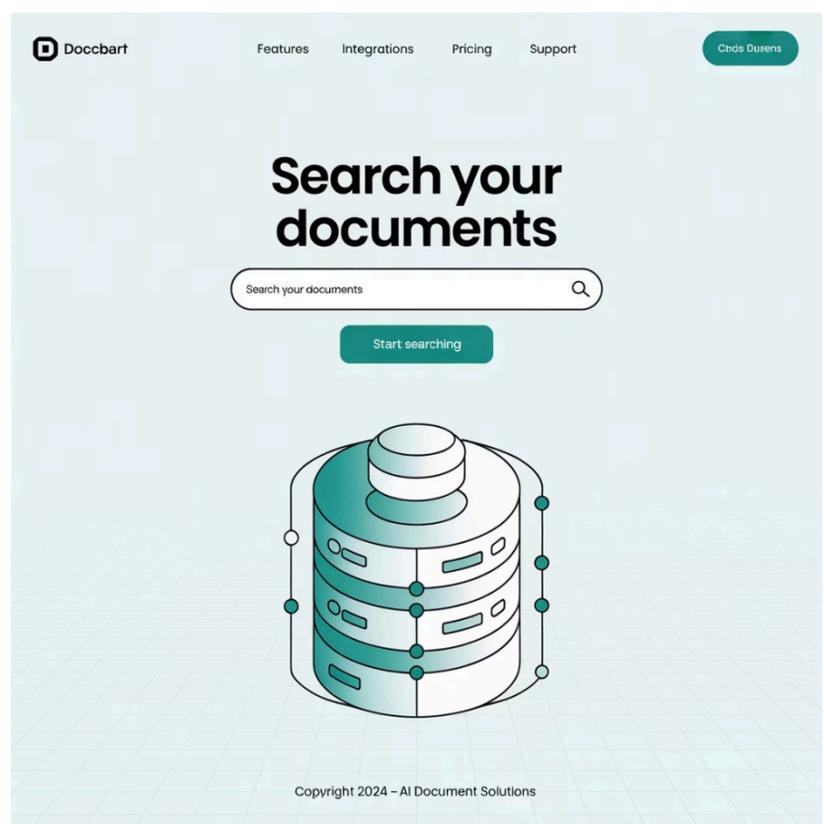
Searching for contract terms, regulations, and KPIs took employees hours. Information was scattered across thousands of documents in Word, Excel, PowerPoint, and PDF formats.

## Solution

Implementation of an LLM-based system with Retrieval-Augmented Generation technology for intelligent document search and analysis.

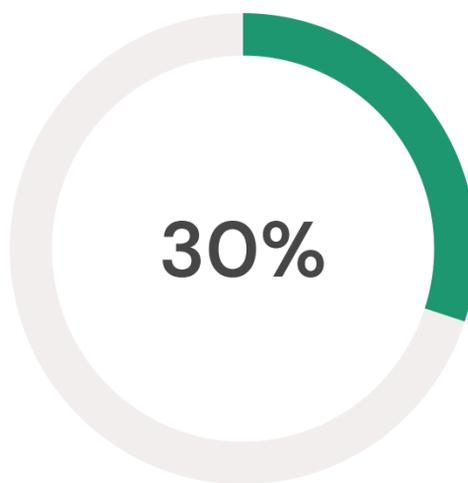
## Solution Architecture

- Parsing and indexing of all document formats
- Vector database for semantic search
- LLM with RAG technology for answer generation
- Web interface with document preview
- Access control system for confidential information



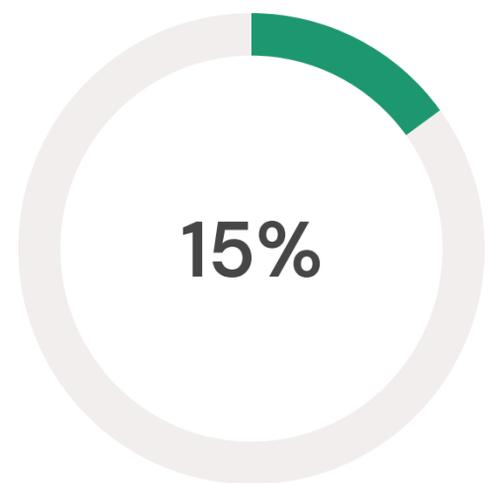
### Answer Accuracy

The system ensures high relevance of search results



### Error Reduction

Automation eliminated human error in searching



### Accelerated Deals

Quick access to information shortened the decision-making cycle

# 5-10 seconds

Answer search time instead of hours of manual work

Technologies: [OpenAI GPT-4](#) • [LangChain](#) • [Milvus](#) • [FastAPI](#) • [Docker](#)

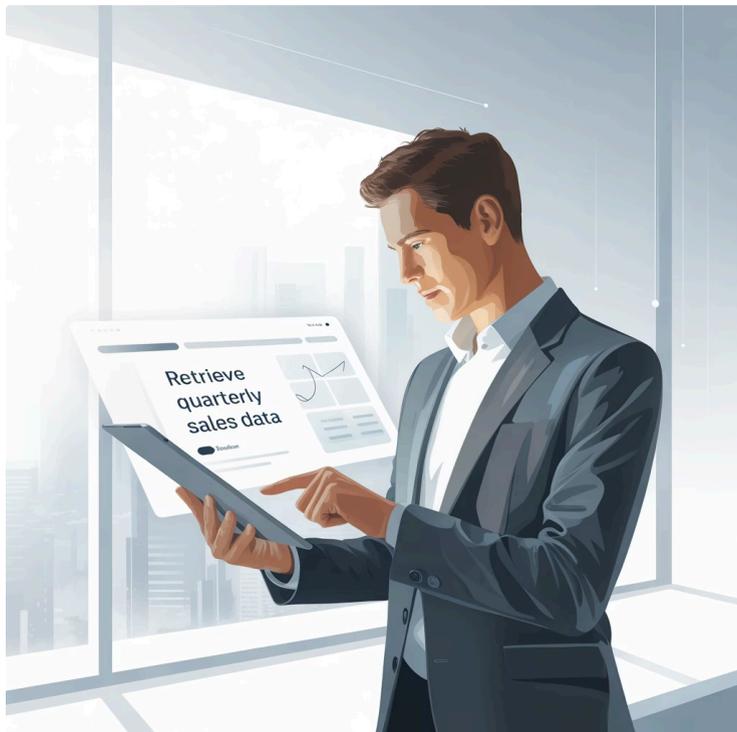
# Text-to-SQL for a Construction Company

## Client

A medium-sized construction company with numerous active projects and a need for operational analytics

## Challenge

Managers couldn't quickly receive project reports — complete reliance on the overloaded IT department created bottlenecks



## Innovative Solution

Development of a text-to-SQL service, enabling managers to retrieve data using natural language without SQL knowledge.

- Generation of SQL queries from natural language
- Automatic validation of query correctness
- Integration with existing BI dashboard
- Security system to prevent unwanted operations



### Natural Language

"Show profit by projects for the last quarter"



### SQL Query

Automatic generation and validation of database queries



### Visualization

Instant display of results in a convenient format

### Report Preparation Time

Reduced from 2-3 days to **5 minutes** — a revolutionary improvement in efficiency

### IT Department Workload

Reduced by **70%**, allowing IT specialists to focus on strategic tasks

### Decision Quality

Managers now make decisions based on **up-to-date data** in real-time

# Voice Assistant for Internal Service



## Voice Interaction

Employees can interact with the system using natural speech to submit requests and obtain information



## Intelligent Processing

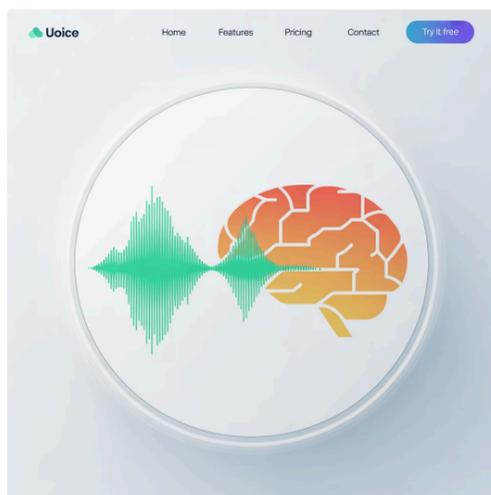
LLM model understands the context of requests and provides relevant answers based on corporate data



## System Integration

Full integration with HRM system and corporate calendar for automating routine processes

*"A company with 500+ employees needed to simplify the processes of submitting requests and finding information through the HR portal"*



## Technical Implementation

The system is built on a modern technology stack to ensure high-quality speech recognition and response generation:

- **Speech-to-Text:** Accurate recognition of Russian speech
- **LLM Processing:** Understanding intent and context
- **Text-to-Speech:** Natural speech synthesis
- **API Integrations:** Connection with corporate systems

### Needs Analysis

Studying existing processes and defining key use cases

1

### System Integration

Connecting to HRM and calendar systems for full automation

2

3

4

### MVP Development

Creating a basic version with core voice interaction features

### Scaling

Deployment for all employees and performance optimization



### 60% Time Savings

Reduction in request processing time thanks to the voice interface



### 200 daily requests

High level of technology adoption by employees



### 40% Satisfaction Increase

Significant improvement in user experience

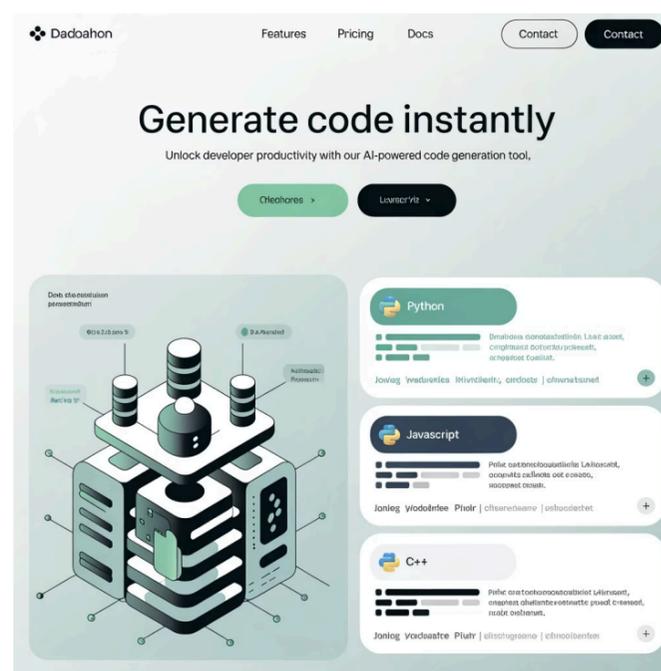
# Self-Hosted CodeLlama for a Startup

 <h3>Data Security</h3> <p>A DevTools startup needed its own code generation model without transferring confidential data to cloud services</p>	 <h3>On-Premise Deployment</h3> <p>Installation and configuration of CodeLlama within the company's closed circuit with full control over the infrastructure</p>	 <h3>Model Customization</h3> <p>Fine-tuning the model on the client's specific codebase to improve the relevance of the generated code</p>
--	---	--

## Solution Architecture

The comprehensive system includes not only the model itself but also the complete infrastructure for its efficient use:

- **CodeLlama Model:** Specialized LLM for code generation
- **Fine-tuning pipeline:** Automatic retraining on new data
- **CI/CD integration:** Embedding into the development process
- **Quality monitoring:** Tracking metrics of the generated code
- **API-interface:** Convenient connection to IDEs and tools

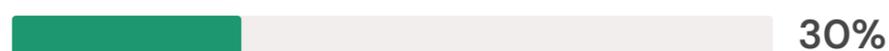


 **Key Advantage:** Full compliance with security requirements while maintaining high-performance code generation



## Time Reduction

For generating boilerplate code



## Accelerated Releases

Thanks to automation of routine tasks

**Technologies:** CodeLlama • PyTorch • FastAPI

# Computer Vision for Defect Detection



## Automated Quality Control

A manufacturing company with a conveyor line needed to reduce defect rates and accelerate product quality inspection



## Computer Vision Technology

The CV model analyzes products in real-time, identifying defects with high accuracy and speed



## MES Integration

The system is integrated with the manufacturing execution system (MES) for automated notifications and quality statistics tracking

## Technical Solution

- **Model Training:** Creation of a dataset with various defect types
- **Real-time Processing:** Image processing at conveyor belt speed
- **Classification:** Automatic determination of defect type and severity
- **Integration:** Connection to MES for production management
- **Notifications:** Instant alerts upon defect detection

01

## Image Capture

High-speed cameras capture product images

02

## AI Analysis

The model processes the image in milliseconds

03

## Decision Making

Automated sorting of good and defective products

### 92% Accuracy

Defect recognition surpasses human control capabilities

### 18% Reduction

In product returns thanks to improved quality control

### 5x Acceleration

Of inspection speed compared to manual control

*The system not only identifies defects but also learns from new data, constantly improving its accuracy*

# Automated Verification of Logistics Documents

**Problem:** Manual verification of thousands of logistics documents in a stone wool insulation company (waybills, CMRs, acts) led to lengthy processing times, numerous errors, and payment delays. An efficient and accurate solution was required to ensure the smooth operation of the logistics chain.

## OCR and Data Extraction

Utilizing Tesseract OCR for accurate text recognition and structured data extraction from scanned documents.

## AI Authenticity Analysis

Developing an ML model based on Qwen-3 for automated verification of signature and seal authenticity.

## Excel Integration

Automated export of verified data to Excel for clients and import of reference information.

**30 sec**

**Document Package Processing Time**

Instead of several days to verify a document package.

**85%**

**Error Reduction**

Thanks to automated verification.

**API**

**Automated Reports**

For instant data exchange and integration with other systems.

**Technologies:** Python • Qwen-3 • Tesseract OCR • Pandas • FastAPI

# Diploma Recognition and Grade Conversion

## Client and Challenge



An American company specializing in verifying diplomas from various countries worldwide. The task is to process diplomas, match subjects to American standards, convert grades into the GPA system, and generate final reports.

## OCR and Structure Analysis



Optical character recognition and document layout analysis technology for extracting information from diplomas, certificates, and transcripts in various formats and languages.

## Classification and Normalization



Automatic classification of academic subjects, their translation into English, and alignment with American educational standards, considering the specifics of different countries.

## GPA Conversion



Automatic conversion of grades from various national systems into the American GPA system, generating a detailed PDF report for the client.



## International Standards

The system works with documents from over 50 countries, taking into account the specifics of national educational systems:

- European ECTS System
- British Classification System
- Russian 5-point System
- Chinese 100-point System
- Indian Percentage System
- And many other regional standards



## Document Upload

Scanning or uploading a diploma in any format



## AI Processing

Extraction and analysis of all academic information



## Conversion

Conversion to the American GPA system

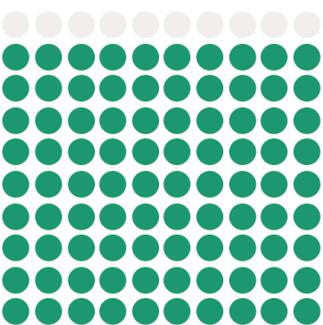


## Report

Generation of a professional PDF document

# 15 minutes

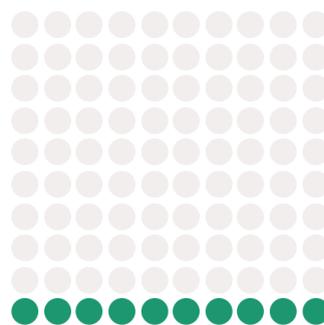
Diploma processing time instead of 2-3 days of manual work



## 90%

### Error Reduction

Automation eliminated human error in processing



## 10x

### Productivity Growth

Increased throughput without expanding staff

- ✔ **Global Impact:** The system processes thousands of documents monthly, helping students and professionals from around the world continue their education and careers in the USA