

Ministry of Education of the Republic of Belarus

Educational Institution
BELARUSIAN STATE UNIVERSITY
OF INFORMATICS AND RADIOELECTRONICS

Faculty of information technologies and control
Department of intelligent information technologies

Admission to the defense:

Head of the IIT Department

_____ D. Shunkevich

EXPLANATORY NOTE

for thesis project
on the topic:

**INTELLIGENT SYSTEM OF FORMING TEXT
DESCRIPTIONS OF INVESTMENT ACTIVITIES**

BSUIR TP 1-40 03 01 02 027 II3

Student

M. Zhirko

Supervisor

D. Shokhalevich

Counselors:

*from the IIT department
in economic affairs*

Y. Krapivin

V. Verniakhovskaya

Standard Examiner

V. Zakharov

Reviewer

N. Mytnik

Minsk 2022

Ministry of Education of the Republic of Belarus

Educational Institution
BELARUSIAN STATE UNIVERSITY
OF INFORMATICS AND RADIOELECTRONICS

Faculty	ITaC	Department	IIT
Specialty	1-40 03 01	Majoring	1-40 03 01 02

APPROVED

_____ Head of Department
«____» _____ 2022 г.

ASSIGNMENT
for the thesis project of
Zhirko Maria

(last name, first name)

1. Project topic: Intelligent system of forming text descriptions of investment activities
approved by order of the university of 10 may 2022 г. № 1133-c

2. Deadline for the student to submit the completed project: 29 may 2022

3. Initial information regarding the project: Development Languages Python; Java; JavaScript.

Purpose of development: investment management with the ability to automate the generation of text.

4. Contents of the explanatory note (list of issues to be developed): _____
Introduction

1 Analysis of approaches to the development of systems of forming text descriptions of investment activities

2 Design of an intelligent system for forming text descriptions of investment activities

3 Development of an intelligent system for the formation of text descriptions of investment activities

4 Feasibility study for the development of an intelligent system of forming text descriptions of investment activities

Conclusion

5. List of graphic material (with the exact indication of the required diagrams): 1

Graphic presentation

2	System user interface
3	Data streams in the system
4	System architecture
5	System usage scenarios
6	Neural network architecture
7	Database model

6. Contents of the feasibility study assignment.

Feasibility study of the development of an intelligent system for the formation of text descriptions of investment activities.

1	Main features
2	Calculation of labor costs
3	Efficiency calculation

Assignment issued: _____ V. Verniakhovskaya

SCHEDULE

п/п	Name of stages of the diploma project	Scope of the stage, %	Deadline for stages	Note
1	Literature selection and study	10	12.02 – 19.04	
2	Study of the subject area, research of analogues	10	20.02 – 19.04	
3	Defining implementation requirements	10	03.04 – 18.04	
4	Designing the system model	15	28.02 – 05.04	
5	System development	30	05.04 – 20.05	
6	Calculation of economic efficiency of the project	5	21.03 – 02.05	
7	Explanatory note formation	10	25.02 – 31.05	
8	Design of the graphical part of the project	10	01.04 – 29.05	

Assignment date: 23.03.2022 г. Supervisor _____ D. Shokhalevich

Assignment accepted _____ M. Zhirko

ABSTRACT

INTELLIGENT SYSTEM OF FORMING TEXT DESCRIPTIONS OF INVESTMENT ACTIVITIES: thesis project/ M. Zhirko. – BSUIR, 2022, – e.n. – 78p., diagrams (posters) – 6 p. A1.

The goal of the thesis project is to provide an automated investment commentary analysis tool for finance professionals. In which comments for individual portfolios are selected based on a modular approach according to performance and contribution, investment strategy and macroeconomic background. Text modules in the system are automatically adjusted by length or content and tailored to the client's interests in the financial sector.

The subject of the research includes basic concepts in the field of finance, principles of their interaction, methods of report generation, workspace organization in the financial system, client-server application development technologies, neural network approach to the generation of highly specialized texts.

The first section of the explanatory note analyzes existing systems for investment activities with similar functionality.

The second section identifies the main entities and the main types of users of this system, design of the conceptual model of data, describes the main functionality of the system, user roles. In this section, the method for automated text generation is selected and the stage of designing the API for communicating with language models is described.

In the third section the analysis of existing technologies, used in the development of such systems is carried out. The basic requirements to the used technologies are defined. Necessary functionality of the system and user interface, which provides the user with interaction with the developed system, are implemented.

The fourth section contains a feasibility study of the developed software product. This feasibility study includes assessment of the economic effect from the use of the system.

The result of the thesis design is an intelligent system for the formation of text descriptions of investment activities.

CONTENTS

Notation list	7
Introduction	8
1 Analysis of approaches to the development of systems of forming text descriptions of investment activities	10
1.1 Analysis of the subject area of investment activities	10
1.1.1 Key concepts and their interaction	10
1.1.2 Types of investments and their characteristics	11
1.1.3 Implementation of investment strategies	13
1.1.4 Conclusion	13
1.2 Software tools for financial strategy management	14
1.2.1 User audience of financial systems	14
1.2.2 The concept of an investment portfolio	14
1.2.3 Structure of holding companies	15
1.2.4 The role of comments and reports in financial systems . .	15
1.2.5 Conclusion	16
1.3 Analysis of existing systems for investment activities	16
1.3.1 Marcus Invest by Goldman Sachs	17
1.3.2 Vanguard Personal Advisor Services	19
1.3.3 Bloomberg Terminal	21
1.3.4 Morningstar	22
1.3.5 J.P. Morgan Wealth Management	24
1.3.6 Comparative characteristics of analogues	26
1.3.7 Defining financial system requirements	27
1.4 Conclusion	27
2 Designing an intelligent system for forming text descriptions of investment activities	29
2.1 General system characteristics	29
2.2 Selection of architectural solutions	30
2.3 Architecture design of the intelligent system of forming text descriptions for investment activities	32
2.4 Description of system functionality	34
2.5 Designing a text processing module	41
2.5.1 Selecting the type of NLP system	42
2.5.2 The text analysis scheme	43
2.5.3 Approaches to the semantic parsing of text	44
2.5.4 Text processing tasks	45
2.5.5 Machine Learning for NLP	46
2.5.6 Comparison of comment generation methods	47

2.6	Designing an API for working with texts	48
2.7	Conceptual data model	50
2.8	Conclusion	52
3	Development of an intelligent system for forming text descriptions of investment activities	53
3.1	Technologies and means of system implementation	53
3.1.1	Technologies and means of implementation of the client part of the system	53
3.1.2	Technologies and means of implementation of the server part of the system	54
3.2	Development of an API for generating text descriptions	54
3.2.1	Paraphrasing text	55
3.2.2	Summarizing text	56
3.3	Database development	58
3.4	Examples of running the application	61
3.5	Conclusion	67
4	Feasibility study for the development of an intelligent system of forming text descriptions of investment activities	68
4.1	Characteristics of the system for forming text descriptions	68
4.2	Calculation of investments in intelligent system development	68
4.3	The economic effect of the development and use of the intelligent system	71
4.4	Calculation of indicators of economic efficiency of the development and use of the intellectual system	72
4.5	Conclusion	72
	Conclusion	73
	Reference list	75

NOTATION LIST

DB – database;
TP – thesis project;
DBMS – database management system;
API – Application Programming Interface;
BOW – Bag of Words;
GPT – Generative Pre-trained Transformer;
HTML – HyperText Markup Language;
IP – Internet Protocol;
JSON – JavaScript Object Notation;
NLP – Natural Language Processing;
ML – Machine Learning;
REST – Representational State Transfer;
TF-IDF – TF – term frequency, IDF – inverse document frequency;
T0 – T5 for zero-shot.

INTRODUCTION

The rise of investment management systems has ushered to a new era in finance. These systems have had a huge impact on the way money managers conduct business. In the not-too-distant past, investment management was done solely by hand. Investment managers used paper documents and ledgers to keep track of their portfolios. This process was time-consuming and error-prone.

With the appearance of investment management systems, the efficiency and simplicity of this activity has increased dramatically. Professionals can now track portfolios electronically using software designed specifically for this purpose. This has led to a number of improvements in the investment management industry.

Perhaps the most significant benefit of investment management systems is that they allow for more efficient portfolio management. With the ability to track all investments electronically, investment managers can quickly determine which investments are performing well and which are not. This information can then be used to make adjustments to the portfolio to improve its overall performance.

Today's systems allow many of the tasks associated with portfolio management to be automated. In the past, investment management professionals had to manually calculate the value of each investment in a portfolio. Nowadays, this task is automated, saving a significant amount of time.

Finally, investment management systems have also made it easier for investment agents to share information. In the past, they often kept their portfolios and records separate from each other. This made it difficult to compare the performance of different investments. Software allows information to be exchanged electronically, making it easier to compare the performance of different investments.

There is now a wide range of investment management software on the market. Each system has its own unique features, but there are also commonalities. When choosing investment management software, it is important to consider the needs of the user.

The target audience of the intelligent system for generating textual descriptions of investment activities are agents. Most investment management agents need access to a system that can improve their performance. Investment management software has the ability to monitor the status of portfolios. This allows portfolio managers to see how investments are performing and make adjustments as needed.

Having a centralized source makes it possible to capitalize on trends and profits, as well as to make better decisions. In the investment industry, mistakes might cost people a lot of money.

Agents have the role of controlling the quality of investments made for

clients. The method of communication and decision-making tool between the client and his financial representative is the portfolio report. The possibility of automated report generation is the main factor that distinguishes the "Intelligent system of forming text descriptions on investment activities" from the analogues on the market.

The purpose of the thesis project is to develop an investment control system with the possibility of automated text generation.

To achieve the goal it is necessary to address the following tasks:

- 1 Examine the subject area of the financial sphere, identify the basic concepts and describe methods of managing them.

- 2 Investigate existing analogues.

- 3 Develop an intelligent system of forming text descriptions of investment activities with the following features:

- control of existing holdings and portfolios;
- data management;
- generation of portfolio reports and comments for holdings;
- saving and uploading data;
- leveraging existing financial strategies.

- 4 Develop APIs for generating reports and comments that perform the tasks of text paraphrasing and summarization.

The diploma project was completed independently, checked in the system "Anti-plagiarism". The percentage of originality corresponds to the norm, established by the Department of IIT. Citations are marked with references to the publications in the "List of references".

1 ANALYSIS OF APPROACHES TO THE DEVELOPMENT OF SYSTEMS OF FORMING TEXT DESCRIPTIONS OF INVESTMENT ACTIVITIES

The history of investment control software can be traced back to the early days of computing, when businesses first began using computers to track and manage their finances.

Early software was designed to simply track financial transactions. Later, it was supplemented with components that could help managers make decisions about where to invest money. This was very important because it allowed businesses to automate investment decisions and free up time for other tasks [1].

Over the years, investment software has become increasingly sophisticated. It began to include features such as risk management and portfolio analysis. This has allowed businesses to make more informed investment decisions and reduce the time it takes to manage finances.

Developing systems to monitor investment activities is a key focus for many businesses and organizations. There are a number of different approaches that can be used to develop an effective system. To implement a relevant system, it is important to understand the needs of the enterprise or organization before choosing the most appropriate one.

The financial industry comprises a set of complex and ever-changing systems. Investment professionals need to stay abreast of the latest news and developments in order to make well-grounded investment decisions. The financial text processing system is designed to be a tool that can help them. The productivity of specialists in this field increases through automation of text processing algorithms.

1.1 Analysis of the subject area of investment activities

Financial sphere is a broad concept that can encompass many different activities and concepts. In its simplest form, the financial sphere can be thought of as the management of money and investments [2]. This can include such activities as banking, investing, and insurance.

1.1.1 Key concepts and their interaction

In any subject area of finance, it is important to understand key concepts and how they relate to each other. In order to make well-informed financial decisions, you must have a solid understanding of concepts in accounting, investment, and financial analysis.

Accounting is the process of recording, classifying and summarizing financial transactions to obtain information needed to make business decisions. **Investment** is the purchase or ownership of assets with the expectation of making a profit. **Financial analysis** is the process of evaluating financial information to make sound investment decisions.

The balance sheet, income statement, and cash flow statement play a key role in the field of accounting. The balance sheet shows a company's assets, liabilities, and equity at a certain point of time. The income statement shows a company's income and expenses for a certain period of time. A cash flow statement shows how a company's cash flow has changed over a period of time [3].

The key concepts in investing are the time value of money, risk and return, and diversification. The time value of money states that money available today is worth more than the same amount of money available in the future. This is because that money can be invested and earn interest. **Risk and return** is the relationship between the amount of risk an investor takes and the potential return an investor expects to receive. **Diversification** — is the practice of investing in different assets in order to reduce the risk of losing money [3].

Ratios and trend analysis are important components of financial analysis. **Ratios** are calculations that show how a company behaves financially. **Trend analysis** is the examination of a company's financial data in order to identify trends.

When considering each company, it is also necessary to consider the amount of its assets, liabilities, and equity.

Assets are all the valuable things that the company owns. These can be tangible assets, such as real estate or equipment, or intangible assets, such as patents or trademarks. Liabilities are any money the company owes to others. This can be money owed to creditors, such as banks, and money owed to employees in the form of wages or benefits. Equity is the difference between a company's assets and liabilities. It is the fraction of a company's assets that is owned by the company's shareholders.

1.1.2 Types of investments and their characteristics

In the broadest sense, investment is the allocation of resources for economic or financial gain. When we talk about investments, we most often think of stocks, bonds, and other securities, but businesses, real estate, and even collectibles can also be considered investments [2].

There are many different types of investments, and each has its own set of risk and return characteristics. Some investments aim for growth, while others provide stability and income. Some are suitable for short-term goals, while others are suitable for long-term goals. Some investments are suitable for first-time

investors, while others are better suited for more experienced investors [4].

The main types of investments and some strategies for using them are reviewed below.

Stocks — are shares of ownership in a state-owned company. When you buy a stock, you become a part owner of the company and qualify for a share of its profits. Stock prices can rise or fall and are affected by many factors, including a company's overall performance, the performance of the stock market in general and economic conditions.

There are two basic types of stock: common stock and preferred stock. Common stock gives you voting rights and capital appreciation, but it also carries more risk than preferred stock. Preferred stock provides stability and income, but generally does not have the same growth potential as common stock.

Bonds are debt tools issued by governments and corporations. When you buy a bond, you lend money to the issuer and are entitled to interest payments. Bonds are generally less risky than stocks, but they have lower yields.

There are many different types of bonds, including government, corporate and municipal bonds. Government bonds are considered the safest type of bond, while corporate bonds are considered riskier. Municipal bonds are a type of bond that are issued by state and local governments.

Mutual funds are investment tools that pool money from different investors and invest it in different securities. Mutual funds are managed by money managers and provide diversification and effective management.

There are many different types of mutual funds, including stock funds, bond funds and index funds. Equity funds invest in a variety of stocks and bond funds invest in a variety of bonds. Index mutual funds track a specific market index, such as the Vanguard High Dividend Yield Index Fund.

Exchange-traded funds or ETFs are a type of investment similar to a mutual fund but traded on a stock exchange. ETFs offer diversification and professional management, but they also have the added advantage of being more liquid than mutual funds.

There are many different types of ETFs, including stock, bond and index funds. Stock ETFs invest in a variety of stocks, and bond ETFs invest in a variety of bonds. Index ETFs track a specific market index, such as The SPDR S&P 500.

Real estate is a tangible asset that can be used for a variety of purposes, including investing, buying a home, and building. Real estate prices can rise or fall and are affected by many factors, including the general economy, interest rates and demographics.

Real estate varies in several types, such as residential, commercial and industrial. Residential real estate includes single-family homes, apartment buildings, and condominiums. Commercial real estate includes office buildings,

retail space, and warehouses. Industrial real estate includes factories, power plants and transportation hubs.

Collectibles are items that are considered valuable to collectors. Collectibles can include a wide range of items, including art, coins, stamps, and sports memorabilia. Collectibles are often bought and sold for profit, but they can also be kept for personal enjoyment.

There are many types of collectibles, each with its own set of risk and reward characteristics. Some collectibles, such as art and coins, may increase in value over time, while others, such as stamps and sports memorabilia, may not. Collectibles also have the added risk of storage and insurance costs.

1.1.3 Implementation of investment strategies

Strategy is an algorithm of actions for investing and withdrawing money. If we talk about investment strategies, each of them has its pros and cons, so it is important to understand the differences before deciding which one best meets your goals [4].

Investment **fixed-income strategies** tend to be low-risk and provide a predictable income stream. However, they may not offer significant capital gains. **Fiduciary investment strategies**, on the other hand, aim to provide both income and capital gains. But they tend to be more volatile, so there is a greater chance of losing money [1].

Balanced investment strategies seek a balance between risk and reward. They usually invest in a combination of stocks and bonds, which protects against losses in any one asset class. **Equity investing strategies**, meanwhile, focus on stocks that have the potential for high returns but also involve higher levels of risk [1].

1.1.4 Conclusion

The key to understanding the financial situation is to analyze the relationship between these concepts. It is important to be able to track changes and how they relate to each other.

There are many different types of investments, such as hedge funds, venture capital and private equity. But the ones listed above — are some of the most common.

Whatever investment strategy you choose, it is important to have a clear idea of your goals and risk tolerance. Based on this, it is possible to choose the strategy and instrument that best meets your stated needs.

1.2 Software tools for financial strategy management

Financial strategy management systems help investors to plan and manage their investment portfolios. There are many programs of this type, and they can be used by both individual investors and investment professionals [1].

When choosing an investment strategy management tool, it is important to consider what type of investor the user is and what their needs are. For example, a novice investor will probably need a program that is easy to use and provides basic functions. On the other hand, investment agents, would need a more sophisticated program with more functions and the ability to track the trends of several portfolios and holdings at once [4].

There is a variety of different features that investment strategy management software can offer. Some of these include the ability to create and manage portfolios, track investment performance, and create and manage investment goals.

1.2.1 User audience of financial systems

An investment system is a set of procedures and related software used to manage an organization's investments. The primary users of the system are usually investment managers and analysts who use the system to make investment decisions, track investment performance, and allocate assets. Other users of investment systems include investment advisors, financial planners, and individual investors [1].

The most important factor in selecting an investment system is its ability to support the organization's investment process. The system should be able to collect and track investment data, provide analysis and reporting tools, and enable transactions. It is important to choose a system that is flexible and can be customized to meet the specific needs of the organization.

Investment systems are becoming increasingly complex as financial markets become more global and volatile. As a result, investment managers have had to become more advanced in their use of investment systems. To make the most informed investment decisions, managers must be able to access data and analytics from a variety of sources. They must also be able to use the system to make transactions quickly and efficiently.

1.2.2 The concept of an investment portfolio

The concept of a portfolio in investment systems refers to a set of investments held by an individual or organization. A portfolio may consist of a variety of assets, including stocks, bonds, mutual funds, and real estate. The composition of a portfolio is determined by the investor's goals, risk tolerance, and timing criteria [3].

Portfolios are often diversified, meaning they contain a mix of assets that are not perfectly correlated. Such diversification can help reduce risk and increase returns. For example, if the stock market falls, a diversified portfolio that includes bonds may not fall as much.

Rebalancing is the process of periodically buying or selling assets in a portfolio to maintain a desired asset allocation. For example, an investor who wants to maintain a 60/40 stock/bond portfolio will rebalance by selling stocks and buying bonds when the stock market performs well and the portfolio becomes too overweight in stocks.

The portfolio concept in investment systems is important for investors to understand. By diversifying their portfolio, investors can reduce risk and potentially increase returns. Rebalancing helps keep the portfolio on track and can be used to buy assets when they are in an undervalued state [3].

1.2.3 Structure of holding companies

A holding company is a type of business entity that owns stock in other companies. A holding company usually does not produce goods or services itself; its purpose is to own stock in other companies to form a corporate group. A holding company can be either a parent company that owns one or more subsidiaries, or a subsidiary company that owns stock in another corporation [2].

The concept of holding companies in investment systems is reasonable. By owning shares in other companies, a holding company can provide stability and growth potential for investors. In addition, holding companies can offer the benefits of diversification because they spread risk across several different companies.

There are several different types of holding companies, each with its own advantages and disadvantages. For example, a holding company can be established as a wholly owned subsidiary, in which case it is subject to the same tax rules as the parent company. Alternatively, the holding company can be established as a separate legal entity, which can provide tax advantages.

1.2.4 The role of comments and reports in financial systems

Investing is a daunting task, especially for novices. Fortunately, there are many resources available to help investors make informed decisions. One of the most important things for investors to do is to stay up-to-date on the latest developments in the market.

This is where monthly commentaries and reports come in. These are documents that provide analysis and recommendations for investments. They usually include information about economic conditions, financial performance of

companies and industry trends. Reports and commentaries are usually written by investment professionals such as investment bankers, portfolio managers or analysts.

While there is no standard format for writing such documents, they usually include an executive summary, investment thesis and recommendations. The executive summary provides an overview of the report and its main conclusions. The investment thesis — is the main argument or rationale for investing in a particular security or sector. The recommendations section provides specific recommendations for buying, selling or holding individual securities [3].

Investment reports and commentary can be useful to both individual and institutional investors. For individual investors, the reports and commentary can provide insight and ideas that can help them make more informed investment decisions. For institutional investors, reports and commentary can be used to support or challenge investment proposals.

Reports are usually published monthly. However, some investment professionals may also publish special reports and commentary as needed. This provides investors with valuable information about what is happening in the market and allows them to make more informed decisions about their portfolios.

Despite the abundance of information available online, not all of it is reliable. That's why it's important to use only reputable sources of information. Monthly comments and reports from reputable firms are a great way to stay informed about the market and your investments.

1.2.5 Conclusion

When it comes to investment software, there are three key terms even though these terms are related, they have different meanings.

A portfolio is a set of investments, and holdings are—the specific securities that make up a portfolio. Reports are—are documents that contain information about a portfolio or holdings.

The main connection between the portfolio, holdings, and reports is that the latter two are used to create the former. In other words, holdings are the building blocks of a portfolio, and reports are used to track and analyze the performance of those holdings.

Although portfolio, holdings, and reports are all important aspects of investment software, they serve different purposes. Understanding the difference between these terms is important for any investor.

1.3 Analysis of existing systems for investment activities

When it comes to investment agents, there are several key features that most, if not all, leading providers offer. These include the ability to track

investments, monitor portfolios and manage client accounts. Other features may include the ability to provide advice and guidance, as well as access to research and market data.

1.3.1 Marcus Invest by Goldman Sachs

Goldman Sachs, one of Wall Street's best-known investment banking names, expanded its consumer banking offerings in 2016 with the creation of Marcus by Goldman Sachs. Marcus, an online bank, offers high-yield savings accounts, high-yield certificates of deposit (CDs) and interest-free personal loans. As of October 2020, Marcus had \$96 billion in [5] deposits.

Marcus, named after one of the bank's Wall Street founders (Marcus Goldman), is positioned as a simple, affordable online banking option that gives people extra control over their personal finances. Goldman Sachs is known as one of Wall Street's most prestigious firms. Marcus has a relatively affordable, "can-do" brand image centered around the goals of helping people achieve financial well-being and manage their money wisely.

Marcus doesn't offer as many types of financial products as other online banks. But the ones it does offer tend to be highly rated. Marcus ranks among the best online savings accounts and has some of the best CD rates. Their personal loans were ranked by J.D. Power as #1 in customer satisfaction with personal loans in 2019. This online bank does not have branches, but it does have a mobile banking app available on the App Store and Google Play [6].

Marcus is an online bank that does not have branches where you can access your account and does not offer an ATM network. However, you can connect your Marcus account to your regular bank account and access mobile banking through the Marcus mobile app.

The Marcus mobile app (picture 1.1) is available for iOS and Android devices. The app allows you to check account balances, view transactions, transfer money to or from other bank accounts, set up recurring deposits, make loan payments to your Marcus personal loan balance and manage your progress toward financial goals. The app is rated 4.8 (out of 5 stars) on the App Store and 3.8 on Google Play.

The Marcus mobile app recently added a "Marcus Insights" feature that allows you to connect external accounts to the Marcus app to track your spending by category, see your monthly cash flow, analyze savings and investment trends and results, and get more visibility to optimize your personal finances.

One possible drawback of the Marcus mobile app is that, unlike some other banks, Marcus can't accept check deposits through the app. If you want to deposit a check into your Marcus savings account, you'll need to mark the check "Deposit Only" and your account number and mail the check to Marcus by U.S. mail.

Features of the tool:

- access to research and analysis from more than 30 providers;
- automatic rebalancing of the portfolio;
- goal-based investing;
- built-in tax optimization;
- real-time investment tracking.

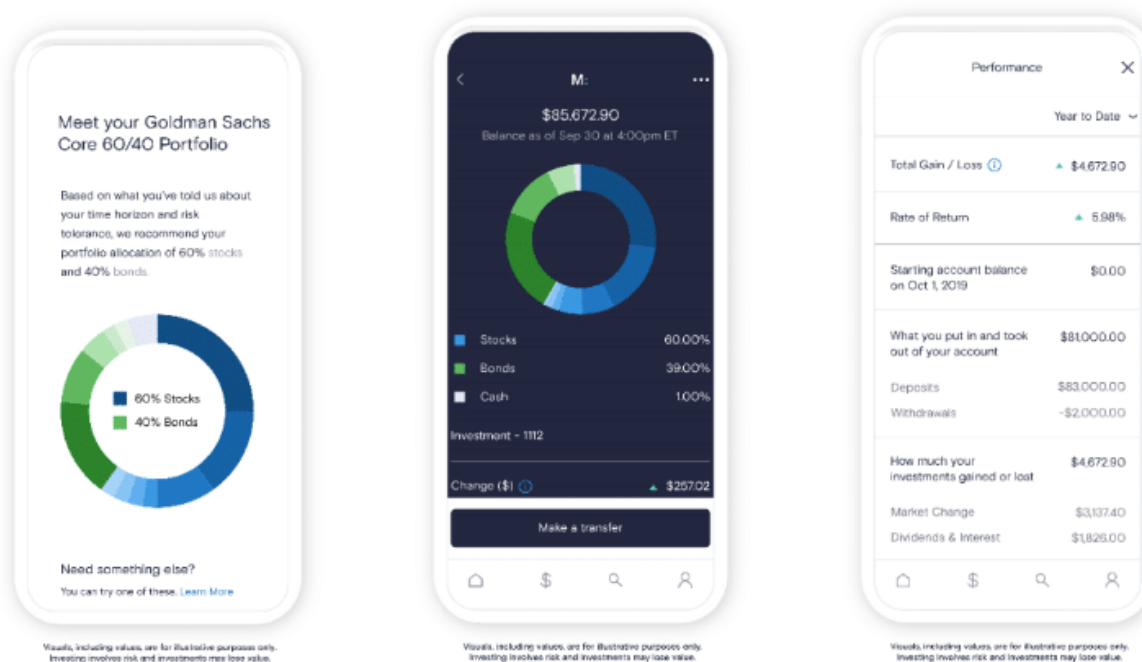


Figure 1.1 – Marcus Invest user interface

Marcus Invest offers three account types: individual investment accounts, joint investment accounts and individual retirement accounts (IRAs). In addition, the \$1,000 minimum account amount and 0.35% advisory commission gives you access to 50 different [6] stock and bond portfolios.

Marcus Invest claims that the advisory fee covers asset allocation, all trading commissions, portfolio monitoring, portfolio rebalancing and ongoing management. The fee also includes ETF expense ratios and ancillary fees [6].

The investment application portfolios are based on three key investment strategies:

- Goldman Sachs Core;
- Goldman Sachs Impact;
- Goldman Sachs Smart Beta.

Goldman Sachs Core portfolios invest your money in market index-tracking ETFs that include a mix of investment-grade bonds, U.S. and international stocks, and real estate securities.

If you're looking for socially responsible investing options, the Goldman Sachs Impact portfolio may be right for you. According to the Marcus Invest website, these portfolios invest in funds that meet environmental, social and governance (ESG) criteria [5].

Finally, Goldman Sachs Smart Beta portfolios seek to provide higher long-term returns with Goldman Sachs ETFs.

1.3.2 Vanguard Personal Advisor Services

Vanguard Personal Advisor Services is a robo-advisor designed with the needs of people who have well-established financial assets and may be approaching retirement. Despite relatively high minimum balance requirements, clients have access to financial advice from live professionals who also manage their investments [7].

Robo-advisors can be divided into two camps: "light" and "heavy". Light robo-advisors are characterized by low account balance requirements, fully automated investing and low commissions. One example is Vanguard Digital Advisor, Vanguard's entry-level robo-advisor.

In contrast, "heavyweight" robo-advisors offer users access to live financial advisors and more customized investment strategies in exchange for a higher minimum balance and potentially higher commissions. Vanguard Personal Advisor Services falls into the latter category.

To open an account with Vanguard Personal Advisor Services, you must have \$50,000 in investable assets. Fortunately, this minimum balance can include funds that are currently already in a Vanguard brokerage account or individual retirement account (IRA), though in either case, there must be at least \$50,000 in Vanguard accounts. The platform charges an annual advisory fee of 0.30% for the first \$5 million in assets under management. The fee is reduced for larger balances [8].

While these requirements make Vanguard Personal Advisor Services unaffordable for most investors just starting out, the competitive annual fee should be very attractive to those looking for a premium robo-advisor with human involvement [8].

Features of the tool:

- access to Vanguard's research and investment analysis;
- individual portfolio management;
- goal-based investing;
- item tracking investments in real time.

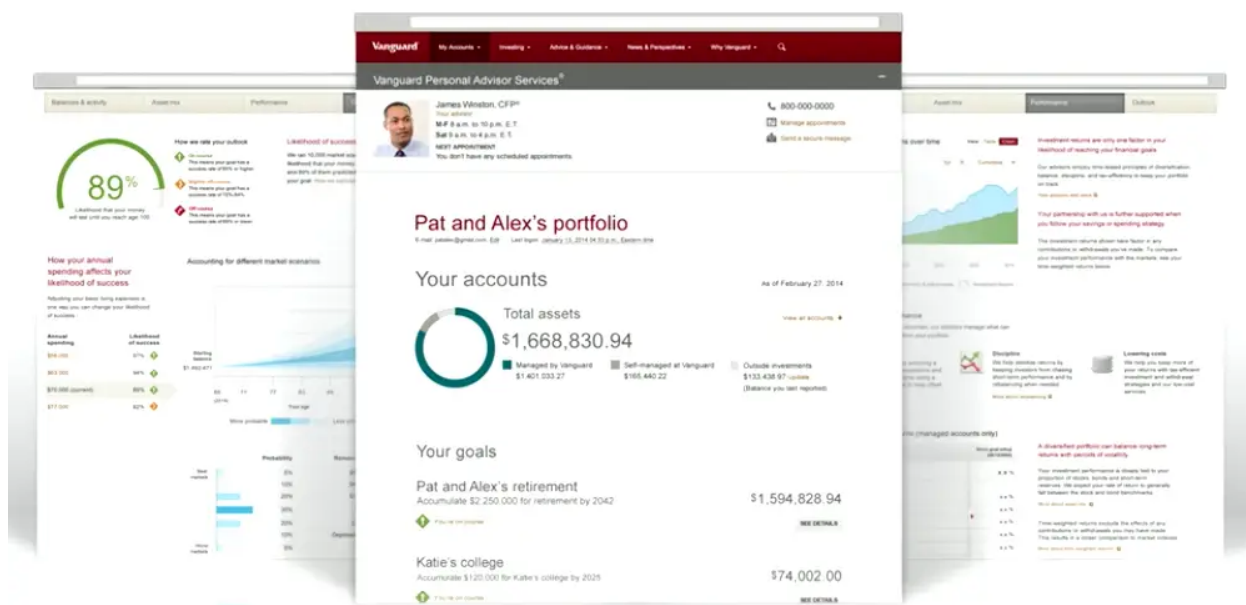


Figure 1.2 – Vanguard Personal Advisor Services user interface

As with most robo-advisors, clients begin their Vanguard Personal Advisor Services journey by completing a questionnaire regarding clients' financial lives.

Clients can and should connect other investment accounts to Vanguard Personal Advisor Services. By connecting their 401(k), Individual Retirement Account (IRA) or 529 account, Vanguard Financial Advisors can get a complete picture of all resources to better guide clients toward their financial goals [8].

After completing the questionnaire and consolidating the client's existing external financial accounts, Vanguard will schedule an appointment for you to meet with one of its financial advisors. The advisor will review the client's current situation and help them clarify their financial goals, using Vanguard technology to build an investment portfolio tailored to their needs.

Those with between \$50,000 and \$500,000 in wealth will be assigned a group of financial advisors, but these clients have no guarantee that they will be able to establish a long-term relationship with any particular advisor. If a client has more than \$500,000 in capital, however, they will be assigned a specific advisor who will handle financial matters.

It is important to note that all Vanguard financial advisors are fiduciaries. They do not receive commissions for their recommendations, which means they have no interest in users choosing one product over another. They will manage a client's portfolio as long as their Vanguard Personal Advisor Services account is active [7].

Vanguard Personal Advisor Services (figure 1.2) are designed for high net worth investors who are beginning to accumulate significant savings in various investment accounts and want a financial advisor to help them plan for the next few decades of their lives. They will receive personalized advice on aspects of their financial present and future and will create a financial plan to manage

their portfolio. The minimum amount of \$50,000, which may be out of reach for some investors, is far less than what it takes for a typical financial advisor to manage client assets.

1.3.3 Bloomberg Terminal

Bloomberg Terminal combines real-time data for every market, breaking news, in-depth research, powerful analytics, communication tools and world-class execution capabilities in one fully integrated solution [9].

The program interface is shown in the figure 1.3.

Awareness, context and analysis make Bloomberg the premier source of information for those who need to know what's happening now and what might happen in the future [10]. Bloomberg Terminal provides unique coverage of markets and securities with information on all asset classes - from fixed income to stocks, foreign exchange, commodities and derivatives - integrated in one place and delivered in real time to your desktop or mobile device.

The company continually develops and applies advanced technology to create best-in-class tools such as custom desktop applications, portfolio monitors, market alerts and charting capabilities that optimize the user's workflow.

The company offers trading solutions for multiple asset classes and sophisticated pre- and post-trade analytics for both the buy and sell side.

Prompt access to news that matters, moves markets and changes the course of business can make a huge difference. News is reported when it happens, from the places where it happens. More than 2,700 news professionals in 120 countries cover companies, markets, economies, politics and governments, enabling Bloomberg to get the information you need, exactly when you need it most [9].

Bloomberg offers a wide range of research offerings, including direct access to vendor-side and independent research from more than 1,500 sources, as well as proprietary analytics-driven research that provides comprehensive insight into industries and their key constituents with unparalleled depth and breadth at the sector, industry and company level [10].

Features of the tool:

- real-time access to financial data and news;
- item financial analysis and modeling tools;
- item built-in chat and text messaging.



Figure 1.3 – Bloomberg Terminal user interface

Bloomberg Terminal is a Windows application that makes it compatible with Excel. Consequently, downloaded data from the tool can be accessed as Excel spreadsheets[9]. Bloomberg can also be accessed from a computer, a smartphone, via the Internet and or through the Bloomberg Anywhere service, which allows the terminal to be used remotely. In this way, investors can view financial information in real time no matter where they are. To work with the terminal, it is necessary to thoroughly study the manual or even ask for help from specialists.

The terminal implements a client-server architecture. The client used by end users to interact with the system is an application that typically connects directly through a router provided by Bloomberg and installed on site.

Multiple instances of the server process run on each server. Using a proprietary form of context switching, the servers keep track of each end-user's state, allowing sequential interactions of a single user to be handled by different server processes.

1.3.4 Morningstar

Morningstar is a Chicago-based investment research company that collects and analyzes [11] data on funds, stocks and overall market conditions. They also provide an extensive line of online products (figure 1.4), software and printed materials for individual investors, financial advisors and institutional [12] clients.

Their analysis covers all corners of the world, including North America, Europe, Australia and Asia. Among Morningstar's many offerings are

comprehensive, one-page reports on mutual funds and exchange-traded funds (ETFs), which are widely used by investors to determine the investment quality of more than 2,000 [12] funds. For example, Motley Fool magazine uses them as its primary source of information.

Morningstar is a respected and trusted source of independent investment analysis for all levels of fund and stock investors, from novice investors to [12] experts. This comprehensive line of products empowers a variety of financial professionals, including individual investors, financial advisors, asset managers, retirement plan providers and institutional investors.

Morningstar's data and research includes real-time [11] information about investment offerings, managed investment products, publicly listed companies and market data. The company's Web site provides free information about individual funds and stocks.

Morningstar can offer a wide range of services, including professional products, managed investment products, and investor products and services.

The line of professional products includes data, research and software to support and strengthen business operations. This includes marketing and reporting solutions, global exchange data, investment planning services and many other products and services. Morningstar also provides rating services for businesses and other financial institutions.

Managed investment products include Morningstar indexes, which provide benchmarking and product creation for financial institutions and professionals; Morningstar managed portfolios, managed advisor accounts and other investment management and indexing services.

Individual investors can take advantage of stock, fund and market analysis, portfolio management resources, and newsletters with investment strategies and other financial and investment materials.

Features of the tool:

- access to research and analysis from more than 30 sources;
- goal-based investing;
- item tracking investments in real time.

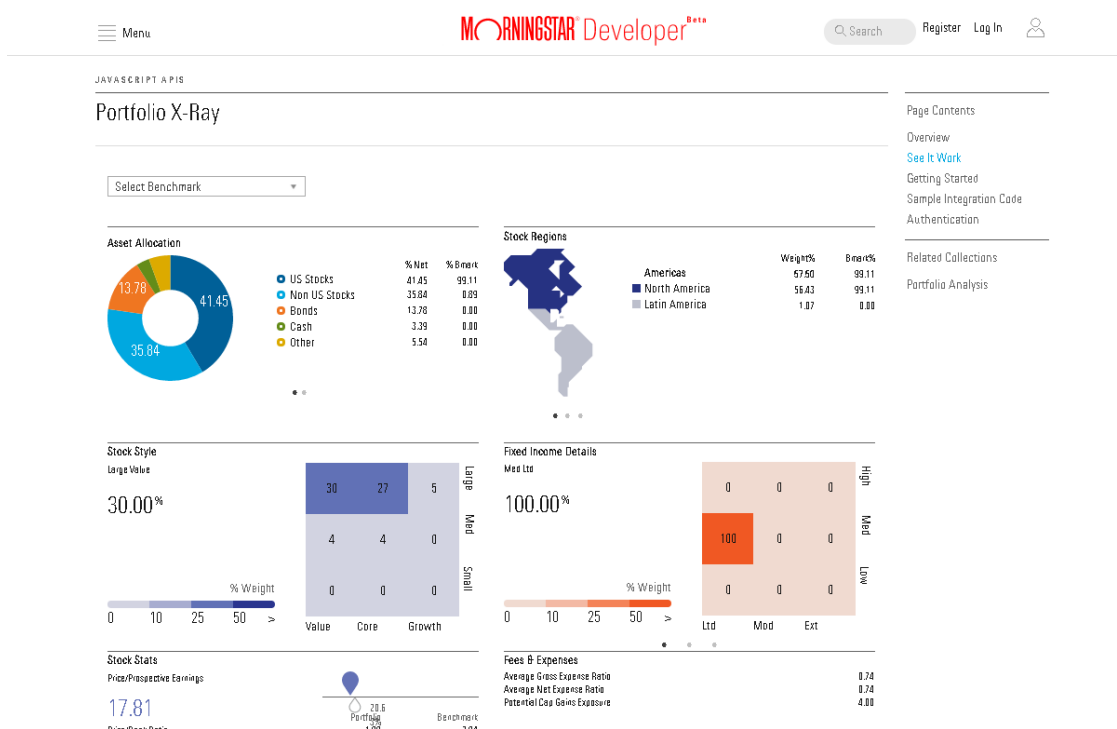


Figure 1.4 – Morningstar user interface

Morningstar educates and empowers investors and other financial professionals to make informed investment and business decisions. Content is available for free or by subscription. Morningstar continues to expand its offerings to further support its strategy to ensure investor success.

1.3.5 J.P. Morgan Wealth Management

J.P. Morgan Wealth Management (figure 1.5) is a division of JPMorgan Chase & Co. that offers investment products and services through J.P. Morgan Securities LLC (JPMS), a registered broker-dealer and investment adviser [13]. Insurance products are provided through Chase Insurance Agency, Inc.

The firm helps individuals and organizations build investment portfolios and manage their wealth. It accepts individual investors, both high net worth and non-high net worth. The firm also serves institutional clients such as retirement and profit-sharing plans, charities, insurance companies and corporations [14].

Among the programs offered, there are various offerings aimed at the wealthy and the insolvent, with specific strategies designed for each [14]. Some are focused on stocks, others on fixed-income securities and different types of funds. Which program is best suited for a client depends on their goals for the future and how much money they have available to invest.

Financial planning is also part of what J.P. Morgan financial advisors can offer. They can cover a wide range of topics such as retirement planning, health care planning, estate planning, social security planning, educational fund planning, wealth planning, charitable gift planning and more [14].

Features of the tool:

- access to research and analytics from J.P. Morgan;
- item individual portfolio management;
- item of real-time investment tracking.

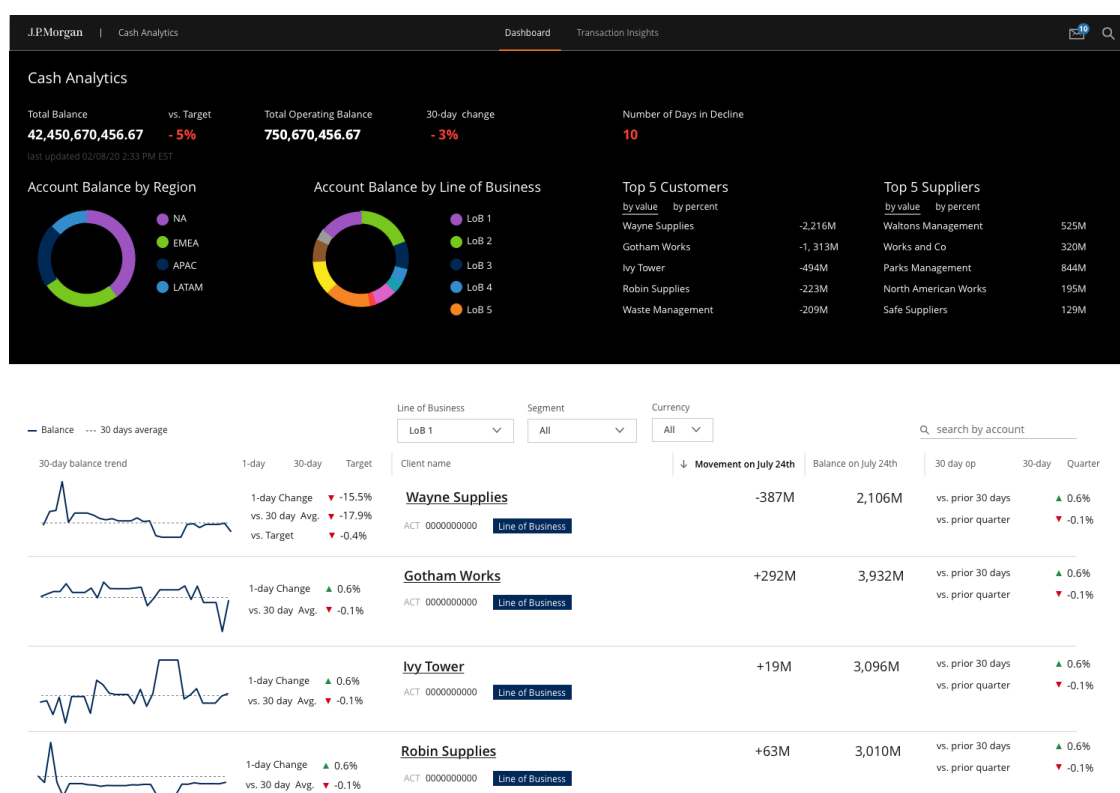


Figure 1.5 – J.P. Morgan Wealth Management user interface

The brokerage company functions as the investment arm of J.P. Morgan Securities, LLC, and offers several low-cost investment options for both active traders and passive [13] investors.

It includes two investment options: J.P. Morgan Self-Directed Investing and J.P. Morgan Automated Investing. J.P. Morgan Self-Directed Investing functions as an online broker and J.P. Morgan Automated Investing provides managed portfolios.

J.P. Morgan online investing is convenient for existing Chase Bank customers who want to try their luck in the stock market. This doesn't mean you have to have a Chase bank account to sign up; the brokerage just makes it easy to connect all your Chase accounts in one place [14].

J.P. Morgan uses both automated investment technology and skilled professionals to maintain portfolio health. This way, clients don't have to worry about making trades or rebalancing the portfolio on their own, as with a self-administered account.

Basically, the platform uses four portfolios: conservative, moderate, growth and aggressive. Each portfolio uses a different set of J.P. Morgan ETFs.

With J.P Morgan, you can trade stocks, ETFs, mutual funds, bonds (that is, buy bonds directly) and other fixed-income assets. The platform tries to distinguish itself from other investment apps by helping investors keep track of their overall portfolio, giving them the ability to choose from a variety of investments.

But it should be noted that although J.P. Morgan Self-Directed Investing is positioned as a "trading" platform, it does not support foreign currencies, cryptocurrencies or futures. Nor can you buy or sell fractional shares using a J.P. Morgan Self-Directed account or trade on margin (margin trading is only offered on full-service accounts).

System user accounts are protected with multi-factor authentication. They are prompted to enter several factors even when logging in from known IP addresses.

Investor accounts are also somewhat protected from theft as long as J.P. Morgan is responsible. Nevertheless, investments in the account are not completely safe. Investments in any brokerage account can lose value and even go to zero. Stock market investors must be willing to risk money if they are going to invest.

1.3.6 Comparative characteristics of analogues

Marcus Invest, Vanguard PAS, Bloomberg Terminal, Morningstar and J.P. Morgan Wealth Management are all software platforms designed for investment agents. All five platforms offer similar features and capabilities, including the ability to research and track investments, create portfolios and make trades. However, each platform has unique features and capabilities that may be more or less tailored to the needs of a particular investment agent.

During the comparison of peers, several areas of focus were highlighted. This was done in order to identify the most sought-after functionality among already existing resources, which can be used when developing your own system.

Comparative characteristics of automated systems for training are shown in the table 1.1.

Table 1.1 – Comparative characteristics of systems for financial activities

Критерий сравнения	Marcus Invest	Vanguard PAS	Bloomberg Terminal	Morningstar	Wealth Management
Possibility to add new tools	–	+	+	–	+
Openness of the platform	+	–	–	+	–
Management considering types of holdings	–	+	–	–	–
Portfolio optimization	+	+	+	+	–
Providing reports	+	+	–	+	+

Marcus Invest — is a cloud-based platform that offers real-time market data, news and analysis. It also provides tools for creating and managing portfolios as well as making trades. Vanguard PAS — is a desktop application that provides comprehensive investment research and analysis as well as portfolio management and trading capabilities. Bloomberg Terminal — is a subscription-based platform that offers real-time market data, news and analysis, as well as comprehensive investment research and trading capabilities. Morningstar — is a subscription-based platform that offers investment research, data and analysis, and portfolio management and trading capabilities. J.P. Morgan Wealth Management — is a platform offering investment research, data, analysis and portfolio management and trading capabilities.

1.3.7 Defining financial system requirements

Investment systems are a useful tool for investors of any level. Such tools must be able to handle large amounts of data, and efficient data processing plays an important role. In addition, the software should provide the user with the ability to easily view and manipulate data.

The software should also be able to generate reports based on the data. These reports should be customizable and exportable to various formats (PDF, Excel, etc.). The software should also be able to email these reports to the appropriate parties.

An important criterion is a user-friendly interface that is easy to navigate. It should be intuitive and not require lengthy training to use. And a scalable system will facilitate its use by businesses of all sizes.

1.4 Conclusion

Analysis of existing financial systems allowed us to conclude that the urgent task is to develop a software solution that would allow the most efficient organization of the working environment for professionals in the financial sphere.

Financial management software is—an important tool for users who are in the stock market. However, it is important to remember that it is only a tool, not a guarantee of success. There are many factors that influence successful investing, and software is only one part of the equation. Quality software will help the user control and manage these factors.

The analysis described the basic concepts found in this type of system. The options for their interaction and the relationship between them have been considered.

The basic requirements and ideas for creating a system have been defined, which are the ability to generate reports on the current state of portfolios in an automated mode, effective compilation of analytics on the performance of holdings, intuitive interface and the availability of basic functionality offered by manufacturers of systems in this area.

2 DESIGNING AN INTELLIGENT SYSTEM FOR FORMING TEXT DESCRIPTIONS OF INVESTMENT ACTIVITIES

The purpose of the TP is to develop and implement a system for generating textual descriptions on investment activities.

The system is a fully automated market commentary analysis tool for individual portfolios based on a robust and efficient approach. All comments for individual portfolios are selected based on a modular approach according to performance and contribution, investment strategy and macroeconomic background. Asset managers can use it as a support or basis for further analysis of individual comments or integrate it as a standard module into individual client communications. The text modules can be automatically adjusted in length or content and adjusted to the client's level of knowledge or interest in the financial field.

2.1 General system characteristics

In accordance with the requirements for the developed software product, and taking into account the implementation of the necessary functions, the following subject area entities are identified:

- System users:
 - Portfolio Manager – the person who decides which financial instruments to buy for the portfolio;
 - Content Manager – a person who prepares the system for next month's reports by writing content;
 - Fund Manager - a person who manages the fund, that is, decides what financial instruments to buy;
 - Administrator – a person who manages the accounts of content managers and portfolio managers;
- Portfolio – a group of investor's financial instruments;
- Holding – a general concept for everything that can be in a portfolio. For example, asset classes, transactions, hedging;
- Asset Class – a group of financial instruments;
- Instrument – any financial instrument that can be bought and put into a portfolio;
- Emerging Market – a region of the world for investment;
- Fund – a financial instrument that invests in other financial instruments;
- Holding performance – the performance of a financial instrument (holding) over a period of time;
- Portfolio Commentary – content generated by the application. It explains the portfolio performance over time.

System requirements:

- portability;
- reliability;
- efficiency;
- flexibility;
- confidentiality.

The main functions of the system under development are:

- user management;
- portfolio and holding management;
- generation of reports and comments;
- ability to export materials containing information about holdings;
- viewing of statistics and progress of portfolio performance;
- availability of translation of text comments into specified natural languages.

Management means functionality such as creation, edition and deletion of data and relevant content. In the system it is necessary to realize the mechanism of differentiation of rights of access to the information according to the allocated roles: the portfolio manager, the content manager, the administrator.

2.2 Selection of architectural solutions

The software product under development architecturally consists of three main components:

- server that performs all necessary functions and provides access to them through the Web API;
- client web application for portfolio managers;
- client web application for content managers.

«**Client – server**» is an architecture in which the network load is shared between service providers, called servers, and service customers, called clients [15].

In fact, the client and the server are software. Usually these programs are located on different computers and communicate with each other over a network using network protocols, but they can also be located on the same machine.

Server programs await requests from client programs and provide them with their resources in the form of data (e.g., file downloads via HTTP, FTP, BitTorrent, streaming media or with databases) or in the form of service functions.

Since a single server program can handle requests from multiple client programs, it is placed on a dedicated computing machine, configured in a special way, usually in conjunction with other server programs, so the performance of this machine must be high. Because of the special role of such a machine in the network, the specifics of its hardware and software, it is also called a server, and

the machines that run client programs, respectively, clients.

The benefits of this architecture are that:

- There is no duplication of the server program code by the client programs;
- Since all computations are performed on the server, the requirements to the client computers are reduced;
- All data is stored on the server, which, as a rule, is much better protected than most clients. It is easier to organize permission control on the server to allow access to data only to clients with corresponding access rights.

Disadvantages

– Server inoperability can make the entire computing network inoperable. An inoperable server should be considered a server whose performance is insufficient to serve all clients, as well as a server that is under repair, maintenance, etc;

- Support of this system requires a separate specialist - a system administrator;
- Expensive equipment.

Client — a local mobile device on the virtual user side that makes a request to the server to be able to provide data or perform a specific group of system actions [16].

At the most basic level of abstraction, a server-oriented application consists of the following architectural layers:

- The core of the application, which includes system components that are not available for user interaction;
- Graphic user interface;
- Reusable components: libraries, visual components and others;
- Environment files;
- Application resources: graphic files, sounds, necessary binary files.

The application core, consists of the following layers:

- Start layer, which defines the workflow, the beginning of program execution;
- The network layer, which provides the mechanism of transport interaction;
- The API layer, which provides a unified command system of interaction between the client and the server;
- Network caching layer, which provides the acceleration of network network caching layer to speed up client-server network communication;
- Layer validation of data obtained from the network;
- Essence layer of data transmitted over the network;
- A data model that provides the interaction of the data entities;
- A local repository.

Server — a very powerful computer or special system hardware which is

designed to solve a certain range of tasks in the process of execution of program codes. It performs service work on client requests, gives users access to certain system resources, stores data or the database [16].

The server is implemented autonomously and does not depend on mobile applications, but exists by itself. The choice to use a server that only provides an API and doesn't take part in the process of creating the user interface gives us the ability to write any client applications, whether they are mobile applications or modules for other systems. For the construction of the server we chose a typical three-tier architecture:

- Data access layer, which provides extraction of information from the database;
- Business logic layer, which provides and controls the operation of the system and performs all operations necessary for the correct execution of the assigned tasks;
- Web interface layer, which provides a way to interact with the server via HTTP protocol, consists of a set of controllers, each of which provides a set of access points for performing specific tasks.

Properties of the three-tier architecture:

- Each layer knows nothing about the subsequent (higher) layers;
- Each layer can interact by control with the immediately preceding (lower) layer through a predefined interface, without knowing anything about the internal structure of all preceding layers;
- Each layer has certain resources, which it either hides from the other layers or provides directly to the succeeding layer (through a specified interface) some of their abstractions.

According to the designed and presented above application structure, there are three basic modules in the architecture of the application: the module providing Web API to access the server, the module with the implementation of business logic, the module of access to the database.

To reduce the connectivity between the components of the system, the principle of control inversion is applied. This principle is a set of recommendations for writing weakly coherent code. Its essence is that each component of the system should be as isolated as possible from the others, without relying on the details of the specific implementation of the other components. In the project being implemented, the implementation of dependencies is done through the constructor.

2.3 Architecture design of the intelligent system of forming text descriptions for investment activities

The architecture of the intelligent system of forming text descriptions of investment activities is based on the principle of modularity of the system,

which allows you to use ready-made blocks of text, which are inserted into the ready-made text, and to change the order of their appearance, as well as to complement them with new blocks.

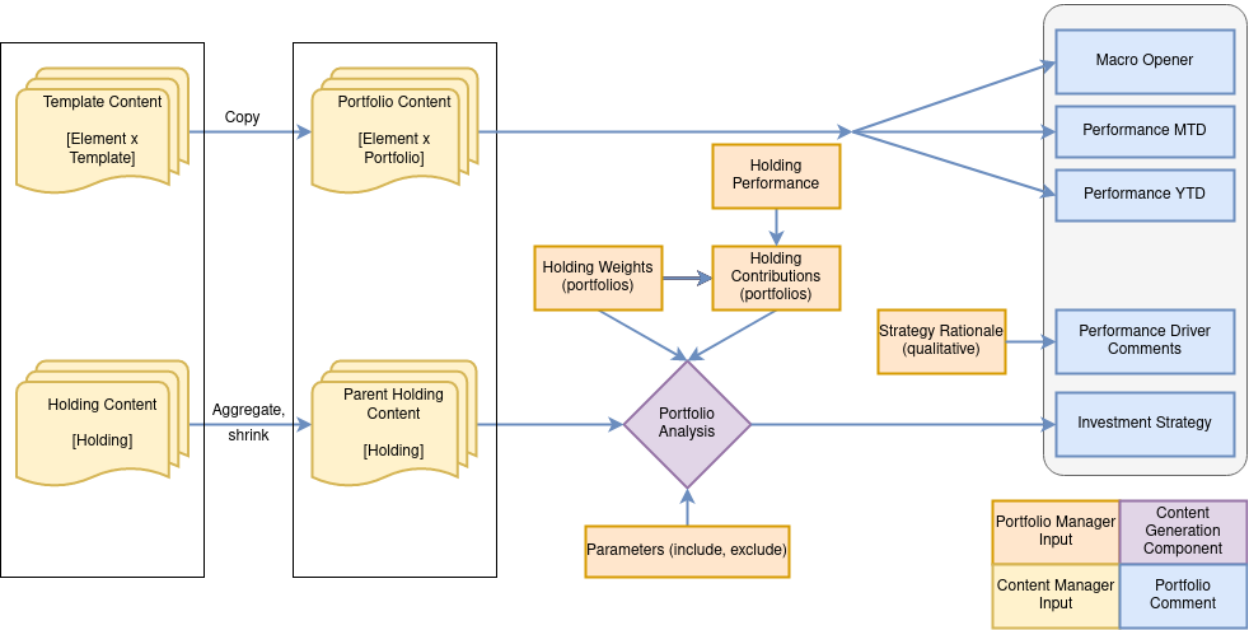


Figure 2.1 – Data flow in the intelligent system of forming text descriptions of investment activities

The data flows in the system are described in the diagram in the figure 2.1.

The system is designed for a clear and concise description of investment activities, highlighting the main stages of the investment process, as well as indicating the key risks and opportunities associated with each stage.

The use of ready-made text blocks allows you to quickly and easily create a text description of investment activities, and the ability to change the order of their appearance and add new blocks allows you to adapt the text to the specific needs of the investor.

The system architecture consists of four components, as shown in the figure 2.2.

- Content DB is responsible for storing data.
- Business Layer is responsible for data processing, API usage and integration with external services as needed.
- Content Management Frontend - data input and configuration for the backend.
- Portfolio Manager Frontend - uploading data and fetching generated content.
- User Manager – user account management for content generation and portfolio management components.

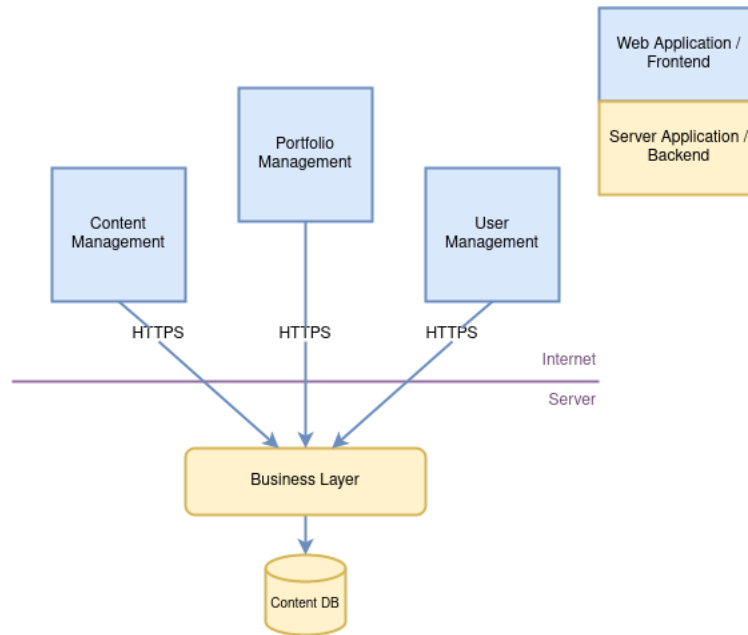


Figure 2.2 – Architecture of the intelligent system for forming text descriptions for investment activities

2.4 Description of system functionality

In order to clarify the functional requirements for the application being developed, as well as the functions assigned to the various user roles, a Use Case diagram was developed.

The essence of this diagram is the following: the designed system is represented as a set of entities or actors interacting with the system with the help of so-called use cases. An actor or actor is any entity interacting with the system from the outside. This can be a person, a technical device, a program, or any other system that can serve as a source of influence on the system being modeled in a way that the developer himself determines. In turn, the use case serves to describe the services that the system provides to the actor. In other words, each use case defines a certain set of actions, performed by the system during the dialog with the actor. It does not say anything about how the interaction of actors with the system will be implemented.

The actors who will interact with the system can be divided into four groups according to their roles:

- portfolio manager;
- content manager;
- fund manager;
- administrator;

Administrator – the person who manages the content manager and portfolio manager accounts. The user interaction with the system is described in the diagram in the figure 2.3.

The administrator is responsible for the following functions:

- System maintenance – ensuring uninterrupted operation of computer equipment, local network and software (regular updates, setup, etc.)
- Management of system data – control of access to the product database, adding accounts to the database server and granting rights, database maintenance (creating backups, compressing versioned databases, updating statistics, rebuilding indexes and updating the database);
- Management of portfolio and holding data – deleting/creating new items as needed;
- User management in the system – creating new user groups, controlling and granting access to groups, deleting user groups.

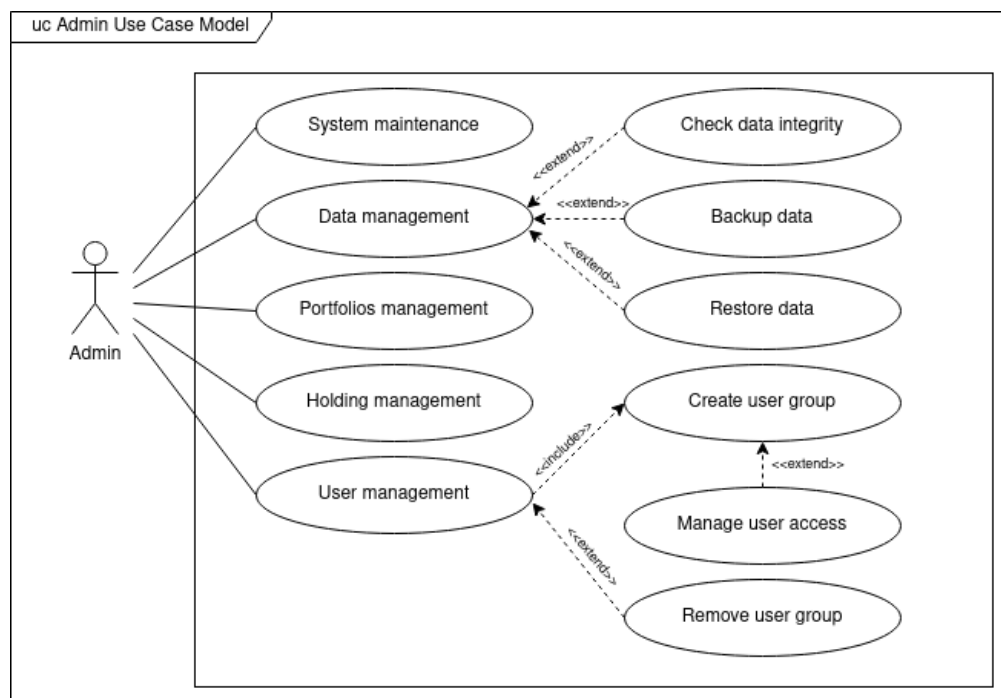


Figure 2.3 – Diagram for scenarios of application usage by an administrator.

The target audience of the system are specialists in the field of finance. Let's consider the envisaged scenarios of actions in the system for this group.

Portfolio manager – a person who decides what financial instruments to buy for the portfolio. The interaction of the user with the system is described in the following diagram in the figure 2.4.

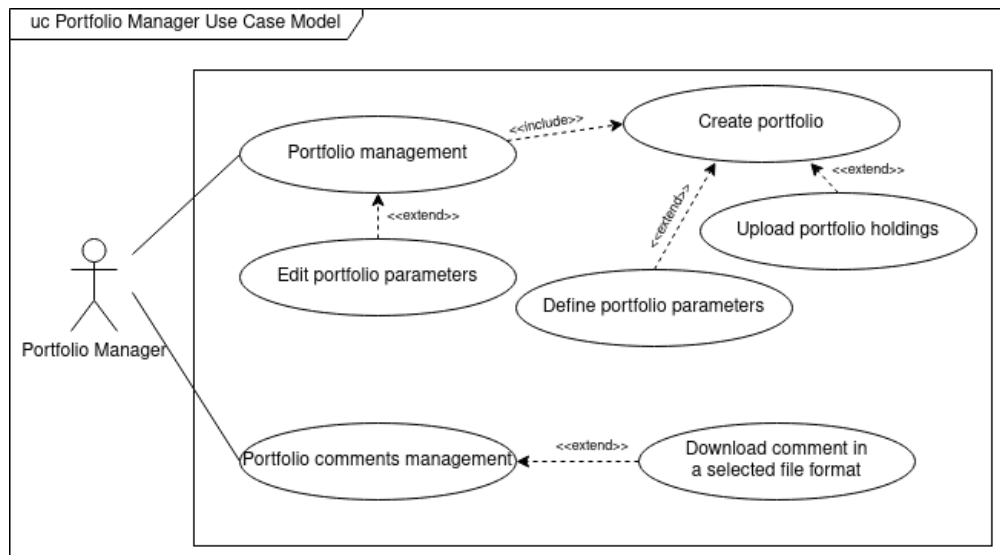


Figure 2.4 – Diagram of scenarios of application usage by a portfolio manager.

The Portfolio manager is responsible for the following functions:

- Portfolio management – creating a portfolio, adding holdings to the portfolio and setting its parameters, editing existing portfolios;
- Portfolio comment management – downloading a file with portfolio comments.

Fund manager - the person who manages the fund, that is, decides which financial instruments to buy. The interaction of the user with the system is described in the diagram in Figure 2.5.

A system function is available to the user «fund manager»:

- Portfolio data management – create a portfolio, define parameters and add holdings to the portfolio, edit parameters of the existing portfolio.

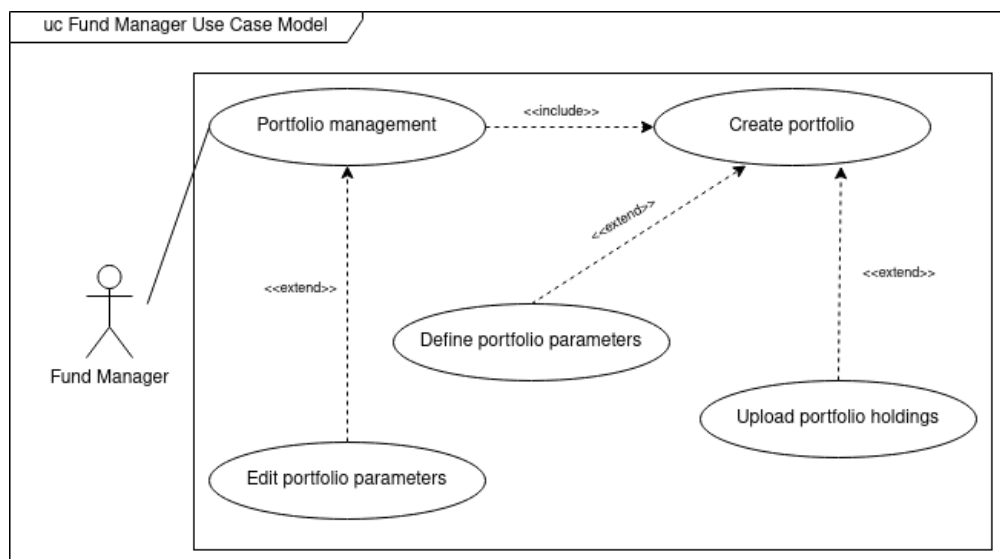


Figure 2.5 – Diagram of the application usage scenarios by the fund manager.

Content Manager prepares the next month's reports for publication,

making comments. The user interaction with the system is described in the diagram in the figure 2.6.

The user of the group «content manager» performs the functions in the system:

- Holding Management – updating the performance data of the holding and its elements;
- Strategy Management – adding strategy descriptions, describing strategy performance over a set period of time, updating performance data for each strategy;
- Holdings Performance Content Preparation – generating and editing content on individual elements of holdings (subsidiaries), consolidating comments on subsidiaries to create a report on the overall performance of the holding and then editing that report.
- Content Preparation by Portfolio – designating performance criteria for companies and holdings to be included in the portfolio report, saving portfolio configurations, and generating, viewing, and editing the final portfolio report.

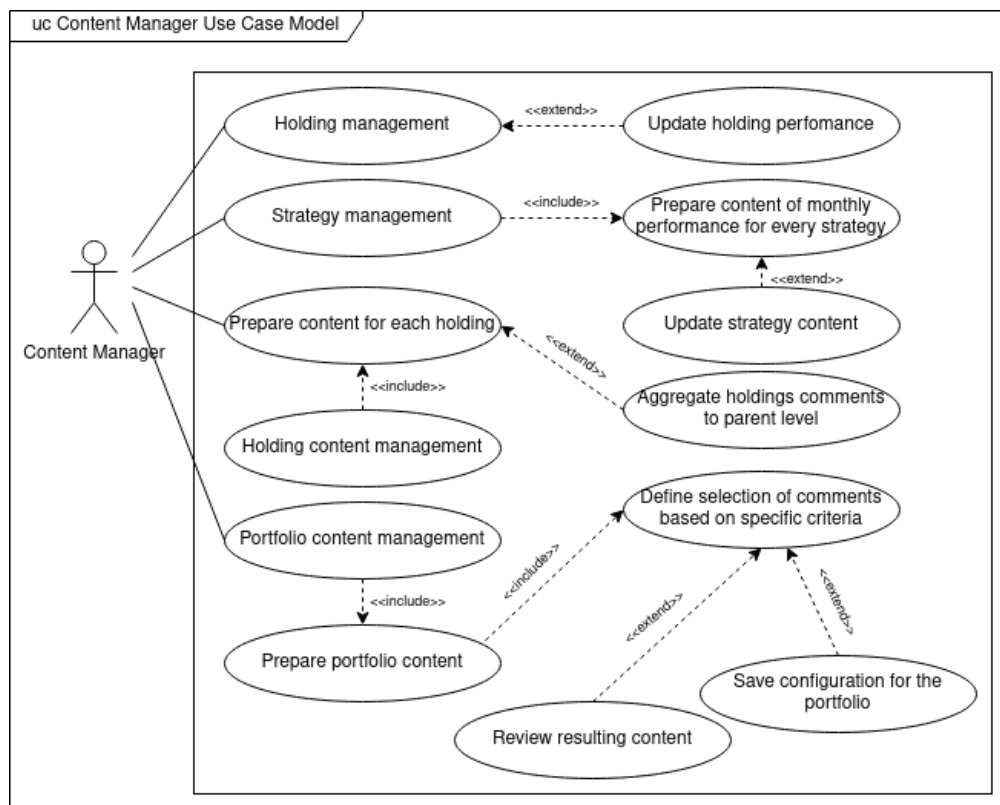


Figure 2.6 – Diagram of application usage scenarios by the content manager.

System user access levels. Each user authorized belongs to a specific group and gets access to the corresponding system component.

The access rights of each user group are marked in the table 2.1.

Table 2.1 – Arrangement of access in the financial system

User role	Portfolio management component	Component of content generation	Component of user management
Administrator	+	+	+
Content Manager	+	+	
Fund manager	+		
Portfolio manager	+	+	

Accordingly, the following scenarios of the software components were formed.

Registration and authorization. The registration and authorization mechanisms are common for all modules of the system. Data about all users, regardless of their role, are stored together.

Registration/restoration process:

- 1 If any customer/employee wants to register, they must contact the administrator and request an account to access the system;
- 2 administrator creates an account using "Administration Console "User Accounts"section;
- 3 Administrator manually sends the client/employee an email with login and temporary password;
- 4 Client must set a new password when logging in for the first time;
- 5 The first account with the administrator role must be originally created by the developers. The details of this account must be kept secret until the first login.

Authorization process:

- 1 The application displays a basic login template (picture 3.5);
- 2 The user fills in the fields in this template;
- 3 The system checks the entered data according to the restrictions;
- 4 The system checks if the user with these credentials exists;
 - item In case the user with such credentials is not found, an error message "User with such login and password is not found" is displayed;
- 5 The system checks if the user's account is active;
 - In case a user with such credentials is found but disabled, the system displays the "Blocked Account Information" form;
- 6 The system checks if the user has one of the allowed roles;
 - In case the user does not have rights to access the system, the "No Rights" information form is displayed;
- 7 The system checks if the password needs to be changed;

- In case you need to change the user's password, the system displays the password setting step.

8 The system authenticates the user and opens the main page of the application.

Logout. Each user can manually log out of the application.

The system should automatically log the user out after N seconds of inactivity (when the user tries to perform a new action). N - configurable on the server side.

Portfolio management component. This component displays all available portfolios and holdings. It provides the ability to manage portfolios, as well as download comments on each item.

Elements available for use when creating a new portfolio:

- 1 Enter the name of the portfolio;
- 2 Choose a suitable strategy (template);
- 3 Selection of the currency of the portfolio;
- 4 Generation of a comment for the portfolio;
- 5 Portfolio download from .csv file;
- 6 Downloading of portfolio data.

Clicking on a portfolio entry in the list opens the portfolio view. Elements available for use when editing a portfolio:

- 1 Enter the name of the portfolio;
- 2 Choose a suitable strategy (template);
- 3 Selecting the currency of the portfolio;
- 4 Generation of a comment for the portfolio;
- 5 Portfolio download from .csv file;
- 6 Downloading of portfolio data.

Elements available for use when deleting a portfolio:

- 1 Select the portfolio to be deleted by name.

Content generation component. The content generation component displays all available portfolio and holding management methods.

The content of the main page of the content generation:

- Holding Level Performance;
- Template Level Performance;
- Holding Level Communications;
- Parent Holding Communications;
- Portfolio Analysis.

The Holding Level Performance section is responsible for displaying and editing the performance of all holdings.

Elements and functions of the Holding Level Performance section:

- 1 Displays information about all elements of holdings
 - date the holding was added;
 - name of the holding company;

- Holdings performance graph during the year;
- Holding's performance factor during the current month;
- value of the market index;
- Bloomberg Ticker;
- type of holding.
- 2 Filter by month the holding or its component is added;
- 3 Holdings search function by name;
- 4 Buttons to save/cancel user action.

The Template Level Performance section is implemented for the purpose of generating content for holding elements.

Elements and functions of the Template Level Performance section:

- 1 Select item;
- 2 Select a pattern (strategy) by the marked item;
- 3 Buttons to go to the next template;
- 4 Button to access the API for text translation;
- 5 Button to access the API to paraphrase the text;
- 6 Table containing:
 - Date of adding/modifying template data;
 - Languages in which the comment is written and the number of characters in each;
 - Template value;
 - Template unique identifier;
- 7 Buttons to save/cancel user action.

The purpose of the Holding Level Communications section is to generate content for the holding.

The elements and functions of the Holding Level Communications section are almost identical to the Template Level Performance section, only the elements and functions change:

- 1 Selection of holding type;
- 2 Selection of the element;
- 3 Buttons to go to the next template;

The Parent Holding Communications section is implemented to combine/reduce the content of the parent holding reports (holding) based on its subsidiary holdings (holding element) and the ability to preview it.

Elements and functions of Parent Holding Communications:

- 1 Holding selection;
- 2 Buttons to go to the next report;
- 3 Button for combining the selected reports into one;
- 4 Button to access the API for text translation;
- 5 Button to access the API for paraphrasing the text;
- 6 Button to access the API for text summarization;
- 7 Buttons to save/cancel user actions.

The Portfolio Analysis section provides the ability to change the way you create portfolio content and view the resulting commentary.

Elements and features of the Portfolio Analysis section:

- 1 The drop-down list shows the portfolio created in the portfolio management component portfolios;
- 2 Buttons to go to the next portfolio;
- 3 Filtering based on asset classes and rows(elements of holdings), which are selected by the buttons on the right;
- 4 Button «Apply» marks all holdings in the «Use» column that have a higher value than the one set;
- 5 Button "Reset" deletes all labels;
- 6 Button # shows the estimated number of characters for the whole investment comment;
- 7 Preview button opens a pop-up window showing the full investment comment for the selected portfolio (read only);
- 8 Click the save button, then this portfolio will become available in the portfolio management component.

User Management Component Features of this component are available only to users of the category «Administrator». Assignments of this module:

- Create, modify and disable user accounts for all modules of the system;
- View a list of all created accounts;
- Navigate through the list (with sorting, filtering, search mechanisms).

Description of the module:

- 1 User with the Admin role can access the Admin Console module;
- 2 Administrator can view a list of all created user accounts;
- 3 Administrator can sort the list of user accounts by specific column;
- 4 The administrator can filter the list of user accounts by specific parameters;
- 5 Administrator can create a new user account;
- 6 The administrator can edit the existing user account;
- 7 Administrator can re-generate new password for the user when editing the user account;
- 8 Administrator can disable an existing active account.

2.5 Designing a text processing module

Creating the project required studying the theoretical part of text processing. Text processing methods, the concept of word embedding, text processing and transformation metrics were studied.

2.5.1 Selecting the type of NLP system

Currently, there are two main approaches: rule-based, or engineering-based, and machine learning-based. Historically, the first is the rule-based approach, which consists in describing the necessary linguistic information in the form of formal rules. The rules are created by linguists or application specialists in a particular problem domain. In the machine learning approach, the source of information is not the rules, but the selected texts of the problem domain.

Among the methods used in the approach, there are supervised learning methods, unsupervised learning methods, and bootstrapping methods. The most commonly used partial learning with a teacher, which builds a mathematical and software model — a machine classifier that knows how to recognize different classes of text units (words, phrases, and other structures) or texts themselves [17].

The classifier is built on a specially marked text corpus (training sample), in which the recognized units (or texts themselves) are assigned labels encoding important features of the recognized units/texts. Learning is the identification of common patterns inherent to texts. Both of the considered approaches have their advantages and disadvantages. Creating rules is labor-intensive and requires rather skilled labor of a linguist. At the same time, the rules are usually declarative and easy to understand, so they are easy to modify and extend.

Machine learning does not require manual labor to compose rules and reduces the development time of systems, but knowledge is needed to select appropriate training methods. In addition, the resulting models (classifiers) have no explicit linguistic interpretation. Machine learning also requires a suitably marked corpus of texts, which is not always possible.

The current trend is to create modular systems of automatic text processing in which different modules can be created under different approaches, for example, **grafematic analysis** module — based on machine learning, and **morphological** module — based on rules. Machine learning is quite often used to process collections of text documents, using a feature-based text model in which features are defined for each document individually. The attributes may be a variety of information characteristics of the text: both linguistic and statistical and structural: for example, the frequency of certain words (or their categories) in the document, the frequency of use of special characters, the ratio of parts of speech of words, the presence of certain syntactic structures or sections of the text, the date of creation, etc. Varieties of feature model are **BOW model** (bag of words — bag of words), in which the text is characterized by a set of its significant words (usually all significant words, more precisely, their lemmas), and a vector text model, in which the specified set is ordered. The vector model is used, for example, in information retrieval, with more complex characteristics, such as TF-IDF for words, rather than words, as attributes. Separately, the

statistical language model (Language Model) is considered, which is built on a representative array of texts by calculating the frequencies of adjacent words (N-grams). Most often bigrams ($N = 2$) and trigrams ($N = 3$) are considered. The model answers the question how likely a given word is to appear if certain words occurred immediately before it. The probabilities are calculated on the basis of collected statistics.

Varieties of models are N-grams of parts of speech, words, text or N-grams of text letters (other models are also possible).

2.5.2 The text analysis scheme

The greatest opportunities and the highest quality of the analysis of texts can be obtained by carrying out its full analysis. However, the difficulties arising in the creation of such an analysis are such that all the theoretical positions developed so far have not yet been implemented in practice. The main problems here are the complexity of the syntactic analysis of the text and the difficulty of creating a full-fledged expert system. To work properly, the text analysis system must be able to analyze a user's text in terms of syntax (the structure of sentences), semantics (the concepts used in the text) and pragmatics (the correctness of the use of the concepts and the purposes for which they are used). The system must then generate its response in an internal representation suitable for logical inference and synthesize its response in natural language. In general, a system that supports full analysis should contain the following modules.

- **Graphematic analysis** — provides identification of syntactic or structural units from the input text, which may be a linear structure containing a single text fragment.

- **Morphological analysis** — provides identification of the normal form from which the word form has been formed, and a set of parameters assigned to this word form.

- **Syntactic analysis** — the most difficult part of text analysis. Here we need to determine the roles of words and their relationship to each other. The result of this step is a set of trees showing these relationships.

Semantic analysis analyzes the text «by meaning». On the one hand, semantic analysis clarifies relationships that post-syntactic analysis has not been able to clarify, since many roles are expressed not only by means of language, but also by the meaning of the word. On the other hand, semantic analysis allows us to filter out some word meanings or even entire parse variants as «semantically incoherent». Semantic analysis stage is the end of the input text analysis. If you need to generate a response, for example, during a dialogue with a user or when translating documents from a foreign language, the syntactic, morphological and graphematic stages are added to the considered stages. Response generation is

inherent in all kinds of dialog systems, some kinds of text abstracting systems, statistical analysis of text, and text generation to varying degrees. The response is selected from some corpus of texts or generated «on the fly» [18].

2.5.3 Approaches to the semantic parsing of text

Vector representation of words. The main idea of the vector semantics model [19] (vector space model, VSM) is—the representation of each text in the collection as a point in a multidimensional space (a vector in vector space). Closely lying points correspond to semantically similar documents. There are already ready-made methods for representing words in vector format, one of the most popular being word2vec. In 2013 Tomasz Mikolov published a paper in which he proposed to use neural networks to train word vectors, but for a smaller dimension: a neural network of the simplest architecture was trained on tuples (word, contexts), the output of which was a vector of 300 elements [20] for each word. Word2vec — a set of algorithms for computing vector representations of words, implements two main architectures — **Continuous Bag of Words (CBOW)** and **Skip-gram**. The **Skip-gram** method (figure 2.7) was used in this work.

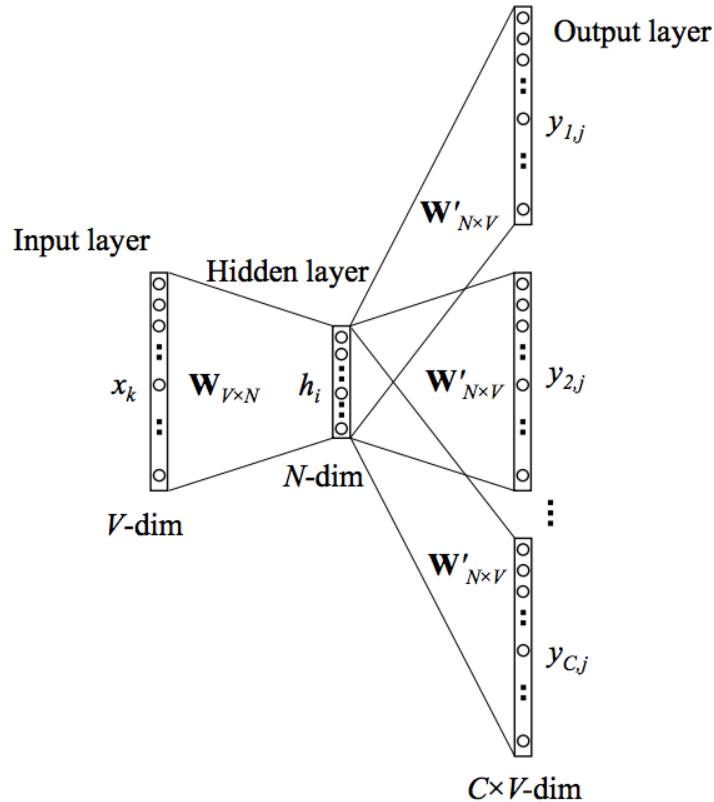


Figure 2.7 – The architecture of the *Skip-gram* method

The input is a corpus of text, and the output is a set of word vectors. Some applications of word2vec technology — identification of semantically related

words, search for typos, evaluation of importance of words in a query.

One of the most popular applications of neural networks — the construction of vectors of words belonging to the field of distributive semantics: it is believed that the meaning of a word can be understood by the meaning of its context, by the surrounding words. Indeed, if we are unfamiliar with a word in a text in a known language, in most cases we can guess its meaning. A mathematical model of a word's meaning is served by word vectors: rows in a large matrix «word-context», constructed from a sufficiently large corpus of texts. Neighboring words, words belonging to the same syntactic or semantic structure with a particular word, etc. can serve as «contexts» for a particular word. The cells of such a matrix can be a record of frequencies (how many times a word occurs in a given context), but more often the Positive Pointwise Mutual Information (PPMI), which shows how non-random was the appearance of a word in a particular context. Such matrices can be successfully used for word clustering or for searching for words that are close to the searched word.

As shown later, word2vec — is nothing more than a factorization of the «word-context» matrix with PPMI weights.

Since the most modern approach is word2vec, then this method can be used with the attention mechanism. In this way, we will get a model capable of giving a response with semantic anchoring to the text.

2.5.4 Text processing tasks

The full-text search and text classification tasks do not require a full morphological analysis of words, but only need to check the fact that the two specified words are actually forms of the same word. This can be done by their lemmatization (reduction to the basic word form) or by stemming, which consists in singling out some unchanged part of the words. However, morphological analysis, lemmatization and stemming are not always able to identify related words, for example, «security» and «protection». The task of determining related words is solved by using special dictionaries-thesauruses, which are oriented graphs, the vertices of which correspond to words, and the arcs correspond to semantically colored relations between words. The proximity of two words is determined on the basis of the shortest path connecting two corresponding vertices of the graph. If it is necessary to take into account the context of words, the task becomes much more complicated and should be referred to semantic text processing. There are automated ways to determine the relatedness of words based on the frequency of their occurrence together or the degree of coincidence of their contexts of use.

When solving practical problems, it is important to divide them into two large groups: processing of individual documents and processing of their arrays.

Text processing for neural network training. The group of tasks

for processing individual documents should be divided into two subgroups: document correction and information extraction. The first one implies that the input and output will be a text document (tasks of error correction, text correction, determination of its structure, abstracting, machine translation). The second subgroup includes tasks related to processing formally represented meaning: fact extraction, natural language queries, organization of natural language interfaces, and generation of correct texts.

Implementation of all tasks of the first subgroup can be based either on rules formulated by experts or on patterns derived from machine learning techniques. The use of a system of rules potentially gives more accurate and predictable results, but implies a high cost of its creation. In turn, the application of machine learning methods is less time-consuming, but requires a large number of high-quality examples. For example, it is possible to find many good translations of English text into other European languages and back, but it is difficult to find examples of translations for rare languages.

The second subset of individual document processing tasks includes information extraction, natural language queries, text generation, and organization of natural language interfaces. All these tasks involve «understanding» texts and searching for referenced facts. The work is based on a system of information extraction rules, each of which specifies a template of syntactic structure and a template of the formed fragment of the formalized representation of information. When the document is processed, the results of the parsing are reviewed and the fragments whose structure corresponds to the templates from the information extraction rules are searched for. Then, according to the "triggered" rules, a part of the words are extracted from the text and converted into a formalized structure.

2.5.5 Machine Learning for NLP

Natural language processing includes speech recognition and generation, classification, knowledge extraction from texts, and other activities aimed at understanding texts in order to populate knowledge bases, form answers to questions, and engage in dialogue. Significant progress in natural language processing technology owes much to machine learning. In machine learning a special place belongs to classification algorithms in tasks related to text processing: spam filtering, sorting documents by subject, selection of named entities. The field of thematic modeling emerged in which documents are considered to be the product of some probabilistic process and consist of a mixture of topics. In computer linguistics, part-of-speech detection has become highly accurate thanks to statistical methods such as hidden Markov chains and maximum entropy models. Neural networks allow to find hidden connections and patterns in texts, but these connections cannot be represented explicitly.

Application of neural networks, first, essentially raises quality of the decision of some standard problems of classification of texts and sequences, secondly, reduces labor input when working directly with texts, thirdly, allows to solve new problems (for example, to create chat-bots). Neural network technologies have radically changed the work with text data. If earlier each text element (letter, word or sentence) had to be described by a set of features of a different nature (morphological, syntactic, semantic, etc.), now in many tasks the need for complex descriptions disappears [21]. Theorists and practitioners of neural network technologies often talk about "representation learning—in a raw text, divided only into words and sentences, a neural network is able to find dependencies and patterns and independently compose a feature space. Unfortunately, a human won't understand anything in this space — during training a neural network assigns to each element of the text a vector consisting of some numbers representing the detected «depth» correlations. The emphasis when working with text shifts from the construction of a subset of features and search for external knowledge bases to the selection of data sources and markup of texts for subsequent training of the neural network, which requires significantly more data compared to standard methods. Neural networks are used in a number of tasks of automatic text processing.

The use of deep learning methods, due to advances in high-performance systems and the emergence of large amounts of data used for training, has eliminated the work of creating the features for machine learning, providing the ability to simultaneously train the extraction of features and training directly to the task itself. New algorithms and approaches, including deep learning, have increased the speed of grammar parsing. In addition, almost all of the leading algorithms and models have become widely available to researchers. Probably the most famous work in deep learning for NLP was the already mentioned word2vec algorithm. It is now common to refer word2vec to distributive semantics rather than deep learning, but the initial impetus for the creation of this model was the application of a neural network. Moreover, it turned out that word2vec vectors serve as a convenient representation of word meaning, which can be fed to the input of deep neural networks used for text classification.

2.5.6 Comparison of comment generation methods

There are several methods to automate the writing of texts: algorithmic and neural network. Each method requires a time commitment from the developer, but the chosen neural network method has its advantages.

For example, a neural network can learn from past data to predict future trends, whereas existing algorithmic methods rely on historical data, and therefore make a less personalized description.

The neural network method relies less on human input and can therefore

make decisions more objectively. This can reduce the risk of human error affecting portfolio performance. Comments that are more personalized to the individual investor matter because it is possible to consider factors such as the investor's risk tolerance and investment goals. The linguistic model can generate more up-to-date commentary because it can be trained on more recent data. In this way, more accurate commentary can also be generated because training is possible on a larger and more diverse set of data.

An important factor is that the neural network method saves the time and resources of the company's employees and therefore allows the creation of quality content for more customers.

2.6 Designing an API for working with texts

For successful automation of text generation in the system it was decided to develop an API for the task of paraphrasing and paraphrasing financial texts. To increase the quality of automated generation, the selected models were further trained on financial texts, so the output is more professional-oriented texts.

API is a set of features and protocols for creating and integrating application software. It is sometimes called a contract between the information provider and the information user, establishing the content required from the consumer (the call) and the content required by the producer (the response) [22].

In other words, if interaction with a computer or system is required to retrieve information or perform a function, the API allows what is required to be communicated to the system so that it can understand and execute the request.

An API can be seen as an intermediary between users or clients and the resources or web services they want to receive. It's also a way to organize the sharing of resources and information while maintaining security, control and authentication – determining who gets access to what.

Another advantage of the API is that the user doesn't have to know the specifics of the caching – how the resource is retrieved or where it comes from.

It was decided to implement a RESTful API for each of the tasks, the architecture of this method is shown in figure 2.8.

REST is a set of architectural constraints, not a protocol or standard. API developers can implement REST in a variety of ways.

When a client request is executed through a RESTful API, it passes a representation of the state of the resource to the requestor or endpoint. This information, or representation, is passed in one of several formats over HTTP: JSON (Javascript Object Notation), HTML, XLT, Python, PHP, or plain text. JSON is the most popular file format because, despite its name, it is language-

independent and easily readable by humans and machines alike.

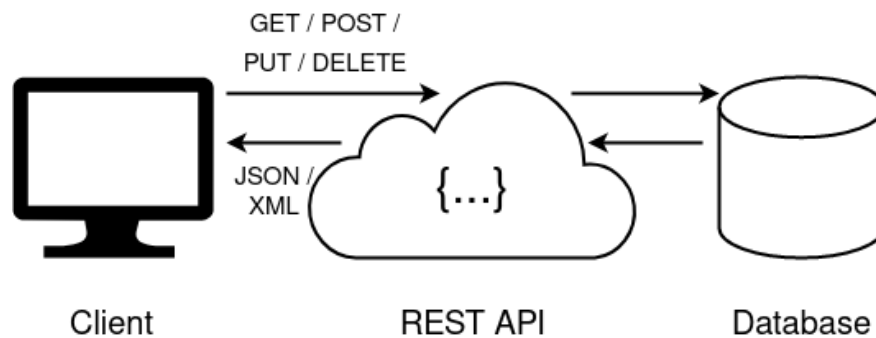


Figure 2.8 – RESTful API model

Although the REST API has certain criteria that it must meet, it is still considered easier to use than a prescribed protocol such as SOAP (Simple Object Access Protocol), which has specific requirements such as XML messaging and built-in security and transactional compliance requirements, making it slower and heavier.

User access to the API is through the use of **portfolio management and content generation** components.

Description of the text translation API. The service DeepL was used for text translation.

The DeepL API provides programmatic access to DeepL machine translation technology, allowing you to use high-quality translation capabilities directly on your own websites and [23] applications. The service uses neural network methods of text processing and shows the best results among existing analogues.

Common use cases include:

- Website translation: Localize websites and reach new markets efficiently and extensively - even in sectors such as e-commerce and news media with a large catalog of dynamic content;
- Company Communications: Integrate DeepL translation technology into your company's systems such as Confluence, SharePoint and Zendesk. Ensure seamless communication between your global teams with maximum data security;
- Create multilingual products: Translate chat conversations to connect users across language barriers in real-time. Localize comments and product reviews at the touch of a button. Make translation one of your differentiating features, however you envision it.

This choice is motivated by the cost and time required for developers to train their own neural network to generate grammatically and syntactically correct text.

Description of the API for paraphrasing texts:

The essence of this service is to provide several variants of synonymous text for further use in a comment.

A model using top-k and top-n sampling. Built on the basis of transformers huggingface [24], TensorFlow and PyTorch. In this case, the transformer model T0 was further trained on the existing investment data.

Algorithms used for paraphrasing texts:

Top-K (Beam Search) is a heuristic search algorithm that explores a graph by expanding the most promising node in a constrained set. Beam search is an optimization of the best-first search method that reduces memory requirements.

Top-P (Nucleus Sampling) - Instead of sampling only the most likely K words, Top-P sampling selects from the smallest possible set of words whose total probability exceeds the probability P. The probability mass is then redistributed among this set of words.

For our source texts, we recommend top-k=50, top-p=0.95, num_return_sentences=3.

Description of an API for retelling texts. Types of summaries

- abstract – this method is rather generative, where phrases and sentences emphasizing key points of the original article are put together to create a clear and concise retelling. The paraphrase need not contain sentences from the original article verbatim. Such summaries tend to be closer to human-crafted summaries, which is what we want to achieve;

- extractive – this approach selects sentences from the original document and inserts them as is in the final retelling. This is easier to do, but often the result is a bit garbled and incoherent.

A model for getting an abstract summary of a report by providing raw text from multiple comments. The GPT-3 [25] model has been used at this point, and its use is motivated by the exceptional text comprehension and text generation capabilities of modern models.

2.7 Conceptual data model

Conceptual design is the construction of a semantic model of the domain, that is, an information model of the highest level of abstraction. Such a model is created without reference to any specific DBMS and data model [26].

Draw up technical requirements, on the basis of which develop specific methods and technical solutions of problems, adopt schematic, algorithmic, software and structural-technological solutions. Determine the users of the system level of abstraction. Such a model is created without focusing on any specific DBMS and data model.

The following basic entities of the data model are highlighted:

- user;

- portfolio;
- holding;
- comment;
- strategy.

Each entity is characterized by a set of attributes. The «User» entity has the following attributes:

- login;
- password;
- password changed;
- username;
- first name;
- last name;
- e-mail;
- role;
- activity in the system;

It is also advisable to create an entity that contains permissible user roles in the system.

Entity «Portfolio», following the described functionality of the system, should contain information about each individual portfolio. Attributes will be required to fully represent the entity:

- unique identifier of the strategy;
- currency;
- name;
- description;
- creator.

«Portfolio» contains some number of holdings, information about the strategy and tools used and their performance, it is acceptable to implement the attributes in a separate entity.

The «Holding» entity describes the element used to manage the portfolio. Its implementation will require the use of the following attributes:

- name;
- unique identifier of the parent holding;
- market index;
- Bloomberg Ticker;
- performance.

«Comment» denotes the text description of the element in the financial system, each element will have its own comment entity with similar attributes:

- unique identifier;
- language;
- comment text.

«Strategy» is an entity that describes a portfolio management tool and has the following set of attributes:

- name;
- description;
- unique identifier.

2.8 Conclusion

At the design stage of the software system, its main components are defined: a server and several client components. Implementation of the server part is proposed in accordance with the three-tier architecture.

The functional requirements of each module of the application are described. Application database conceptual model is built, attributes for each entity are defined.

The motivation for the creation of modules for automated text generation is explained, the requirements and methods of implementation are described.

3 DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR FORMING TEXT DESCRIPTIONS OF INVESTMENT ACTIVITIES

3.1 Technologies and means of system implementation

The purpose of this work is to develop an intelligent system for investment activities using modern information technology.

The system should provide the user with the ability to quickly obtain and process information about the current state of the market, the financial state of the holding, the prospects of its development, as well as make reports on the performance of each asset.

For this purpose, the system should have access to databases of different types of information, which should be constantly updated. In addition, the system should have an artificial intelligence module that will allow it to analyze this information and make reports.

It is also necessary to develop a user interface that would be convenient and intuitive to use.

The system should be implemented based on modern information technology, including artificial intelligence and big data.

3.1.1 Technologies and means of implementation of the client part of the system

The main task of the client is to allow the user to interact with the system, that is, to send requests and receive responses.

The system allows you to use different clients for different purposes. For example, the system can be accessed from the command line, through a web browser or from a mobile device.

The client part of the investment system is implemented mainly in JavaScript and HTML.

It is responsible for communication with the server, the display of investment data for the user and the ability of the user to interact with the system.

The main libraries and tools used are TypeScript, React and MobX.

TypeScript — is a programming language for web development based on JavaScript [27]. It makes code clearer and more robust, adds static typing (variables are bound to specific data types), and can be compiled into JavaScript.

TypeScript helps reduce the time it takes to identify and fix bugs that are sometimes hard to find in a dynamic JavaScript environment. With TypeScript,

you can write more understandable and readable code that describes the subject area as much as possible. In this way, the architecture becomes more pronounced.

React is a JavaScript library for creating user interfaces. React can be used to develop single-page and mobile [28] applications. Its goal is to provide high speed, simplicity and scalability. As a library for developing user interfaces, React is often used with other libraries such as MobX, Redux and GraphQL.

MobX is a library for simple application state management. It is a flexible tool for managing application state. MobX is concise. It allows you to work with the state of components in a synchronous style. In addition, MobX is convenient and does not interfere with TypeScript types. You can describe state as separate classes or directly inside a React component.

As a set of user interface components used Ant Design - React UI library, which provides a set of high-quality components.

The client part of the investment system was designed to be easily extensible. The main libraries and tools used were those that are open source projects and a large community of developers.

3.1.2 Technologies and means of implementation of the server part of the system

The server part of the system is developed using Java, and the main framework is Spring Boot. The team also uses OpenAPI to develop the system more efficiently. Using Java and Spring Boot will allow the team to develop a robust and scalable system. In addition, using OpenAPI will allow the team to develop the system faster and more efficiently [29].

Java is a versatile and popular programming language that is well suited for developing the server side of an investment system. The Spring Boot framework provides a convenient way to develop and deploy Java applications [30]. The OpenAPI specification ensures that the investment system API is well defined and easy to use.

When developing the server part of the investment system, it is important to consider the needs of the system users. The system should be able to handle a large number of simultaneous requests, as well as allow users to easily access their investment data. In addition, the system should be able to generate reports on the effectiveness of the investment portfolio.

In order to meet these requirements, the server of the investment system must be designed to be scalable and reliable. Using Java and the Spring Boot framework will help ensure that the system meets these goals.

3.2 Development of an API for generating text descriptions

The API was implemented using Python and the FastAPI library.

Python is a versatile modern high-level programming language whose advantages include high performance of software solutions and structured, well readable code [31].

Python’s syntax is as lightweight as possible, allowing you to learn it in a relatively short time. The kernel has a very user-friendly structure, and the wide list of built-in libraries allows for an impressive set of useful functions and features.

Python is an interpreted programming language which is not compiled. Thus, before you run it is a plain text file. Accordingly, you can program on almost on all platforms, and the language itself is logical and well designed.

FastAPI is a web framework for developing RESTful APIs in Python. FastAPI is based on Pydantic and type hints for validating, serializing and deserializing data as well as automatically generating OpenAPI documents. It fully supports asynchronous programming and can work with Uvicorn and Gunicorn [32].

FastAPI is faster than Flask because it is built on ASGI (Asynchronous Server Gateway Interface), so it supports parallelism/asynchronous code. This is achieved by declaring endpoints with `async def` syntax. A good feature of FastAPI is documentation.

Two Python APIs have been implemented to handle comments. Each is written for a specific task: paraphrasing and forming a paraphrase.

3.2.1 Paraphrasing text

The program must change the syntax, structure, word/phrase order, and uses synonyms when appropriate. The rewriting algorithm digs into the meaning of the text and looks for other ways to convey the same meaning in other forms. At the same time, it corrects grammatical errors.

The tool uses machine learning and NLP to understand the syntactic, lexical, and textual nature of the language so that the text can be rewritten with the correct context. The API looks at different parts of speech, including nouns, pronouns, verbs, adjectives, adverbs, prepositions, and conjunctions.

The API is implemented based on T0. T0 (T5 for zero-shot) demonstrates a generalization of zero-shot tasks to English sentences in natural language. It is a series of encoder-decoder models trained on a large set of different tasks given in natural language queries [33]. Multiple data sets with controlled data in English are converted into sentences, each containing several templates with different wording. These sentence datasets allow us to compare the model’s ability to perform completely unfamiliar tasks specified in natural language. An example of how the paraphrasing API works is shown in the figure 3.1.

```
[ 'We maintain an underweight in bonds despite the recent developments.',
  'We maintain an underweight in bonds despite the recent developments which have reduced risk.',
  'We maintain an underweight in bonds despite the recent developments that have reduced risk.',
  'We maintain an underweight in bonds despite the recent developments which have reduced the risk.',
  'We maintain an underweight in bonds despite the recent developments which reduced risk.' ]
```

Figure 3.1 – Example of paraphrasing a sentence "Despite the recent developments which have reduced risk to some degree we maintain an underweight in equities versus bonds." with a language model T0

Neural networks have been shown to effectively paraphrase text. In a subsequent evaluation of the performance of a neural network in a paraphrasing task when it is not provided with training data, it is found that the network is able to learn from the data it is provided and generalize it to new data, achieving high performance.

3.2.2 Summarizing text

The API compiles an abstract paraphrase of several texts. Abstract paraphrasing is a method based on a generative approach in which phrases and sentences emphasizing key points of the original article are gathered together to create a clear and concise summary. Summaries do not necessarily contain sentences from the original article verbatim. Such summaries tend to be closer to human-crafted summaries, which is what we want to achieve.

For implementation we use the GPT model. On language tasks (answering questions, reading comprehension, summarizing, translating) it works well without pre-learning.

GPT-3 is a large transformer-based language model with 175 billion parameters, trained on a dataset of 570 GB texts [34]. GPT-3 is trained to predict the next word, given all previous words in the text (figure 3.3). GPT-3 outperforms other language models trained on specific domains (such as Wikipedia, news, or books) without having to use specific datasets for training. For language tasks such as answering questions, reading comprehension, paraphrasing, and translation, GPT-3 starts learning from the source text without using task-specific training data.

GPT-3 zero-shot tasks allow you to use the understanding of your task's underlying model, provided through instructions or headings in the prompt, without reference to examples. The model is guided only by the headings and instructions and uses them to develop your understanding of what you think is right. After all, most GPT-3 tasks are somewhat tentative. What counts as the correct answer and what the model thinks is correct and what the client thinks is correct may differ somewhat. In a retelling task, it may mean an emphasis on certain key words, themes, or phrases. It may be of a certain length or contain specific proper nouns.

The main advantage of a zero-shot paraphrase is the ease with which the system can be constructed. If you understand the construction of the task at a high level, you can achieve a fairly high level of accuracy in a mass of different usage scenarios. After all, the level of system that is needed to solve a particular problem depends entirely on what the problem is and what range of data needs to be covered.

This is what the result of the API will look like (figure 3.2), written using GPT-3.

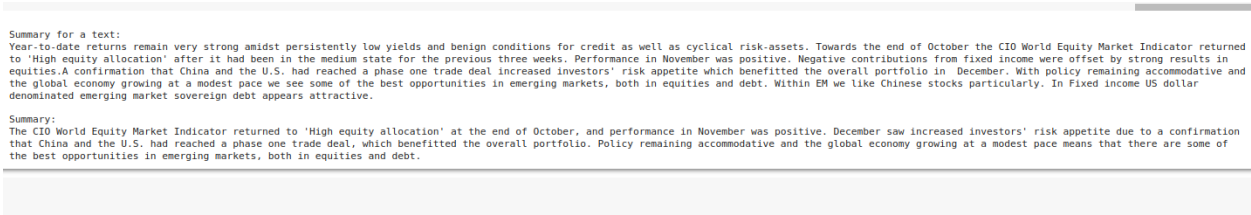


Figure 3.2 – Example of a summarization module task

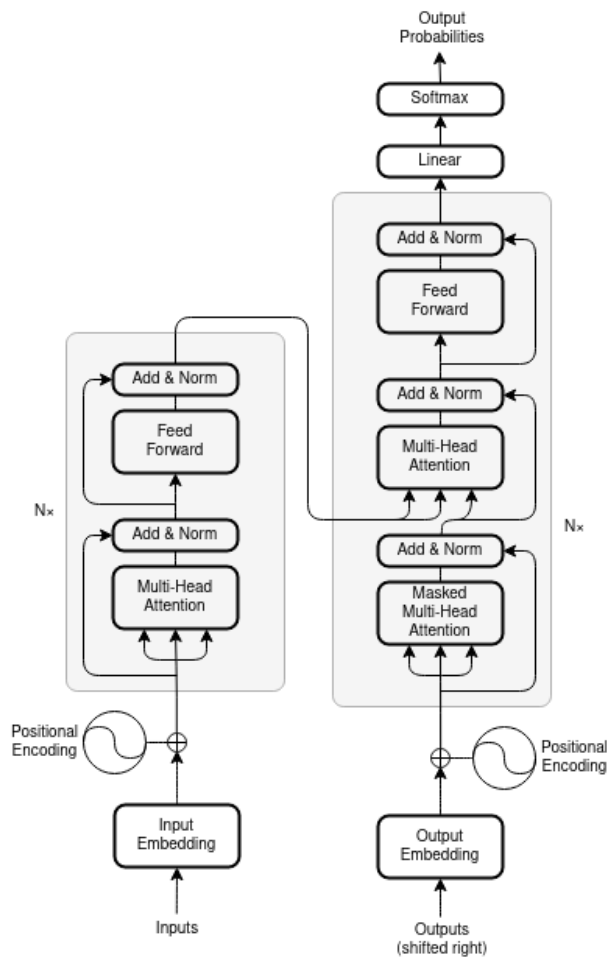


Figure 3.3 – Transformers architecture

3.3 Database development

Database development is based on a preliminary analysis of the subject area, on the main and auxiliary functions identified, as well as a description of the environment of each function. Using this information gives the developer a rationale for the use of the necessary entities in the modeling and presentation of the necessary attributes in them, organizing the correct relationships between entities. Selected approaches in database design: classical (from documents) and functional-object (from functions and objects), — implement a unified database design methodology, determining, depending on the chosen approach, the key element of the analysis - a document or object, respectively.

The development of the classical approach to relational database design is based on the fact that the main element of the user's work in the subject area is a document containing the necessary information that is appropriate to store in a database with subsequent presentation in electronic and paper documents. This approach practically does not take into account the peculiarities of working with data within the individual functions of the user of the information system, for which the database is developed. Developers remember about the functional usefulness of the database only at the time of development of data processing procedures and implementation of the software application.

Thus, the classical approach allows the creation of an indivisible information system, which is good for functional systems that include closely related tasks solved by the user. Creation of complex multifunctional systems led to duplication of information, excessive complexity of the database structure, large time response of the system to user needs.

As a result of solving these problems the functional-object approach to database development was formulated, which, to a greater extent, is supported by object-oriented database management systems, but can also be implemented within classical relational DBMS [35].

The basis of the functional-object approach to database design is the awareness of the needs of the users of the system to perform individual functional tasks, which, as bricks, make up a complex functional environment of the user's activity. This approach, allocating separate functions of the subject area and dividing them into main and auxiliary, allows the developer to approach the design process, considering the functions as independent blocks (black box) and combining them into a structure called "information system".

For data storage we used DBMS — PostgreSQL.

PostgreSQL is an open source database, which has a good reputation for its reliability, flexibility and support of open technical standards. Unlike other RDBMSs (relational database management systems), PostgreSQL supports both non-relational and relational data types.

Originally developed in 1986 as an extension of INGRES (an open-source

relational SQL database project started in the early 1970s), POSTGRES, now known as PostgreSQL, was the brainchild of Michael Stonebraker, a computer science professor at Berkeley. SQL support was added to the project in 1994, and PostgreSQL [36] appeared shortly thereafter.

Maintaining dynamic database systems is critical in today's digital space, especially given the speed of new technologies. PostgreSQL is extensible and versatile, so it can quickly support a variety of specialized use cases through a powerful ecosystem of extensions that cover a wide range of data types, from time series to geospatial analytics.

Its feature-rich and affordable design makes PostgreSQL a "one-stop" solution for many businesses looking for cost-effective and efficient ways to improve their database management systems. Designed as an open source database solution, PostgreSQL is completely free of licensing restrictions, vendor lock-in potential, or the risk of over-deployment. Expert developers and commercial enterprises who understand the limitations of traditional database systems actively support PostgreSQL.

In large database systems where data authentication and read/write speed are important, PostgreSQL is hard to beat. PostgreSQL supports many performance optimizations typically found only in proprietary database technologies, such as support for geospatial data and unlimited concurrency.

When multiple users access data simultaneously, traditional database systems typically block access to records to avoid read/write conflicts. PostgreSQL effectively manages concurrency by using MVCC (Multiversion Concurrency Control). In practice, this means that reads do not block writes and writes do not block reads.

PostgreSQL is one of the most flexible databases for developers because of its compatibility and support for multiple programming languages. Popular programming languages such as Python, JavaScript, C/C++, Ruby and others have advanced support for PostgreSQL, allowing developers to perform database tasks in the language they speak without causing system conflicts.

Enterprises need to maintain continuous operation in the event of disasters. They need a resilient solution that ensures that production databases are always available to both customers and developers. PostgreSQL can be configured to provide high availability services using asynchronous or synchronous replication methods on multiple servers.

Deploying open source database management technology gives enterprises unique benefits, including lower cost, greater flexibility and innovation not always available with proprietary database solutions.

Based on the designed conceptual data model, a database is developed. Figure 3.4 shows the main entities and relationships of the database model.

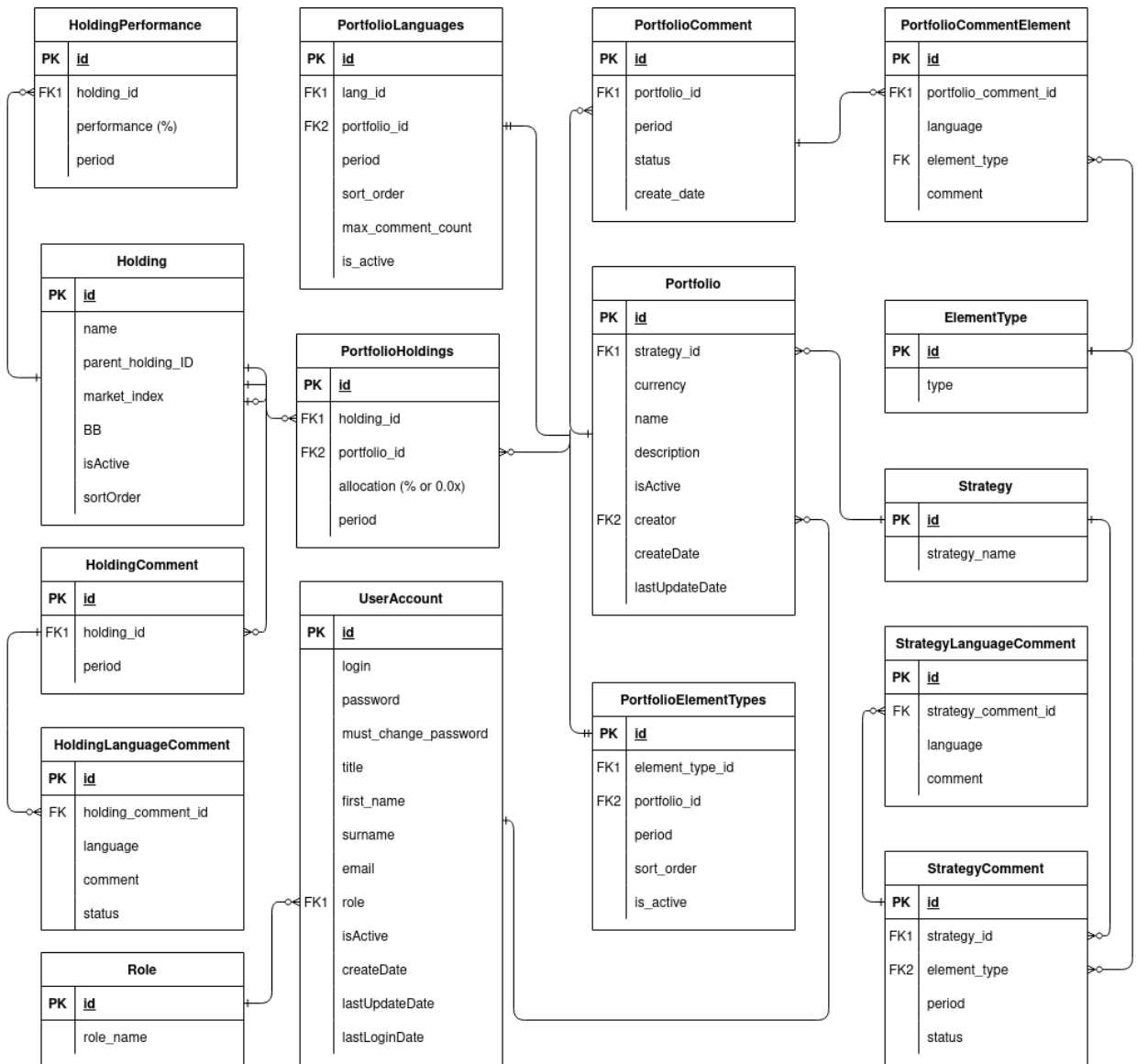
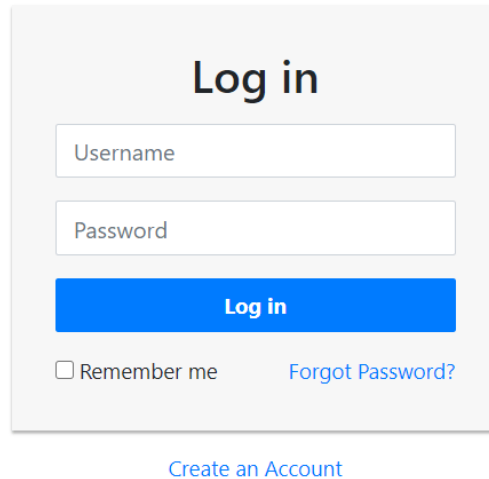


Figure 3.4 – The main entities and relationships of the database model

3.4 Examples of running the application

Intelligent system of formation of text descriptions of investment activities. The user opens the login form component (picture 3.5), enters his login and password, and clicks the login button. The system then checks the user's credentials and, if they are correct, displays the main content of the site. If the credentials are incorrect, the component displays an error message.



The image shows a login form with a light gray background. At the top, the text "Log in" is centered in a bold, dark font. Below this, there are two input fields: "Username" and "Password", each with a light gray border. Under the "Password" field is a solid blue button with the text "Log in" in white. Below the button, there is a checkbox labeled "Remember me" and a blue link "Forgot Password?". At the bottom of the form, centered, is a blue link "Create an Account".

Figure 3.5 – The main login form.

After successful authorization, provided the user is in the group accessing the content generation and portfolio management components, the system displays the main screen (Figure 3.6), which contains the following components:

- Holding Level Performance;
- Template Level Performance;
- Holding Level Communications;
- Parent Holding Communications;
- Portfolio Analysis.



Figure 3.6 – Main application page

When navigating to the Holdings Performance View page, the user will see the following screen (Figure 3.7). It allows you to view information about holdings and their attributes. Filtering of holdings by date of creation, search by name is available. The user can also add a new holding, provided that he is a portfolio manager or fund manager.

Holding Level Performance

filter by month

search

holding name

YY/MM	holding name	year chart	monthly	market index	BB	holding type
20/01	NYSE:DASH		-10.30%	BB Barclays Swiss Franc Aggregate AA or H	LEGATRUU	AssestClass
19/02	ETR:DHER		-13.26%	JP Morgan Asia Credit Index JACI IG	BDSYTRUU	AssestClass
22/03	AMS:TKWY		-8.03%	ICE BofA Euro High Yield Constrained	LD06TRUU	AssestClass
21/06	HKG:3690		-4.68%	MSCI Emerging Markets	I18785	AssestClass
22/07	GOOG		-1.57%	BB Global Aggregate - Corporates	LG30TRUU	AssestClass
20/02	NASDAQ:FB		-2.54%	UBS Bloomberg CMCI Ex Agri Ex Livestock	HUC0	AssestClass
19/01	NASDAQ:AAPL		-2.87%	Rogers Int Commodity Agricultural Index T	JACINGTR	AssestClass
20/11	HKG:0700		-4.69%	World Inflation-Linked Securities Index	JCMDCOMP	AssestClass
21/07	AMS:VWRL		-2.56%	JP Morgan CEMBI Diversified Index USD u	UCBIGLBL	AssestClass
22/08	AMS:PRX		-5.08%	FTSE World Government Bond Index WGE	UCBIUS	AssestClass
21/02	JSE:NPN		-2.23%	S&P Risk Parity Index 10% Target Volatility	MXWD	AssestClass
22/05	AMS:ADYEN		-1.00%	Invesco Global Listed Private Equity ETF. F	HFRIFO	AssestClass
20/08	GOOG		-1.57%	SBI Domestic Government (TR)	CMXALCTR	AssestClass

save

cancel

Figure 3.7 – Holdings performance page

When switching from the home page to the template creation screen, the user can select an item and a template to create, view or edit its description. In this screen (picture 3.8), the user has the ability to access the API on the formation of text comments through the buttons rewrite and translate. Editing is available after the user selects the text fragment, which he wants to change.

Template Level Communications

Element

Choose ▾

Template

Choose ▾

translate

rewrite

<

>

YY/MM	20/01
Eng	327
De	315
Templ	7
ID	1205

We are neutral on USD duration and tactically over USD-denominated EM sovereign bonds against high interest rates. On the currency side we currently express a preference for GBP, JPY as well as a basket of EM currencies vs. the market currencies.

Wir stehen der USD-Duration neutral gegenüber und haben eine taktische Übergewichtung von auf USD lautenden Staatsanleihen der Schwellenländer gegenüber Hochzinsanleihen. Auf der Währungsseite bevorzugen wir derzeit USD, GBP, JPY sowie einen von Schwellenländerwährungen gegenüber den Währungen der Industrieländer.

Figure 3.8 – Page for creating templates

When selecting the Holding Level Communications component on the main page (picture 3.9), the user has the functionality to generate comments on companies in holdings. To generate a report on a subsidiary in the holding, he needs to select the type of holding and the name of the subsidiary. After that, the text field for entering a comment will be available to him. Automated editing is done through the translate/rewrite buttons. The user also has the ability to upload a comment for display on the web page or download the comment in the desired format. To control the user's actions, save/cancel buttons have been added.

Holding Level Communications

Holding Type

AssetClass

Holding

HGBUSD

translate

rewrite

<

>

YY/MM	20/01
Eng	327
De	315
Templ	7
ID	1205

We are neutral on USD duration and tactically over USD-denominated EM sovereign bonds against high interest rate risk. On the currency side we currently express a preference for GBP, JPY as well as a basket of EM currencies vs. the market currencies.

Wir stehen der USD-Duration neutral gegenüber und haben eine taktische Übergewichtung von auf USD lautenden Staatsanleihen der Schwellenländer gegenüber Hochzinsanleihen. Auf der Währungsseite bevorzugen wir derzeit USD, GBP, JPY sowie einen von Schwellenländerwährungen gegenüber den Währungen der Industrieländer.

Figure 3.9 – Comment generation page for companies in holdings

Unlike the previous screens, when the user selects the Parent Holding Communications component on the main page. To create comments on the parent holding, information about all its subsidiary holdings is displayed (3.10). Their contents are then combined into a single text file (concatenate button) after which the user can generate a summarized report and edit it.

Parent Holding Communications

Holding

HGBUSD

summarize

translate

rewrite

<

>

We are neutral on USD duration and tactically over USD-denominated EM sovereign bonds against high interest rate risk. On the currency side we currently express a preference for GBP, JPY as well as a basket of EM currencies vs. the market currencies.

Wir stehen der USD-Duration neutral gegenüber und haben eine taktische Übergewichtung von auf USD lautenden Staatsanleihen der Schwellenländer gegenüber Hochzinsanleihen. Auf der Währungsseite bevorzugen wir derzeit USD, GBP, JPY sowie einen von Schwellenländerwährungen gegenüber den Währungen der Industrieländer.

concatenate

Figure 3.10 – Report generation page by holdings

On the Portfolio Analysis page (Figure 3.11), the user can view the portfolios created in the Portfolio Management component. Filter the content

based on the selected components. The preview button opens a pop-up window that displays a full investment commentary based on all of the user's selected components.

Portfolio Analysis

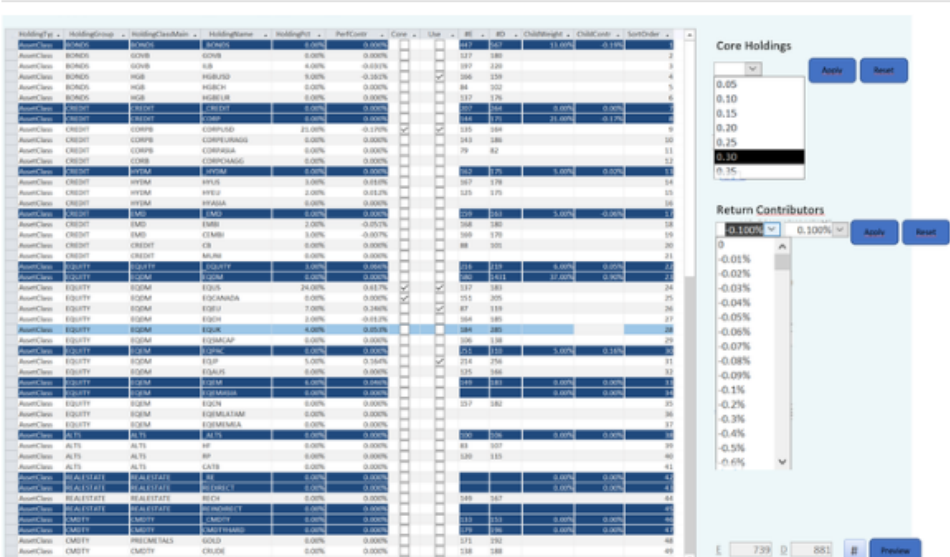


Figure 3.11 – Portfolio performance analysis page

The software component in 3.12 is displayed to users of the Content Manager and Portfolio Manager groups. This page displays all available portfolios.



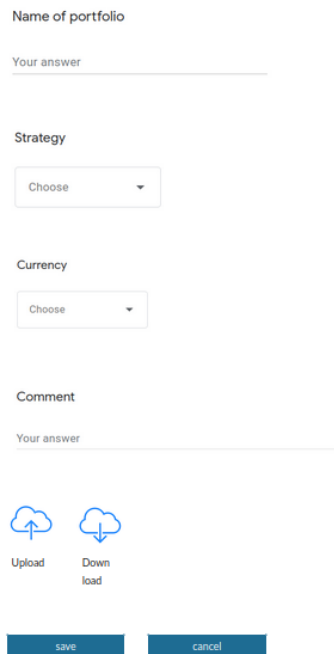
Existing Portfolios

portfolio	strategy	currency	comment	upload portfolio	download comment
Equity Japan	Balanced	CHF	initiated 02.03.2021		
High Dividend Switzerland	Equity	CHF			
Equity Switzerland	Balanced	CHF	initiated 07.09.2022		
Corporate Bonds EU	Fixed Income	EUR			

Figure 3.12 – View available portfolios page

If you want to create a new portfolio, the user fills in the 3.13 component fields. The fields will be filled in automatically when the data is loaded from a file. You can also download the contents of the newly created portfolio.

Create Portfolio



Name of portfolio

Your answer

Strategy

Choose

Currency

Choose

Comment

Your answer

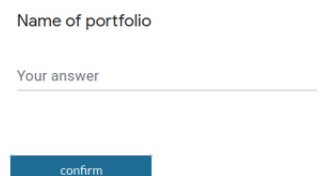
Upload Down load

save cancel

Figure 3.13 – New portfolio creation page

If you want to delete a portfolio, the user goes to the 3.14 screen and selects the portfolio you would like to delete.

Remove Portfolio



Name of portfolio

Your answer

confirm

Figure 3.14 – Delete existing portfolio page

The holdings view page provides information about the holdings (figure 3.15). If you want to view the subsidiaries, the user selects the parent holding for which he would like to display information.

Holdings

holding name	company name	year chart
NYSE:DASH	DoorDash Inc	
ETR:HER	Delivery Hero SE	
AMS:TKWY	Just Eat Takeaway.com NV	
HKG:3690	Meituan	
GOOG	Alphabet Inc Class C	
NASDAQ:FB	Meta Platforms Inc	
NASDAQ:AAPL	Apple Inc	
HKG:0700	Tencent Holdings Ltd	
AMS:VWRL	Vanguard FTSE All-World UCITS ETF USD Dis	
AMS:PRX	Prosus NV	
JSE:NPN	Naspers Limited	
AMS:ADYEN	Adyen NV	
GOOG	Alphabet Inc Class C	

Figure 3.15 – Page for browsing holdings

3.5 Conclusion

The intellectual system of forming text descriptions of investment activities, which meets all the requirements and solves the whole list of tasks set for it, was developed, namely:

- User Management;
- Portfolio and Holding Management;
- Generation of reports and comments;
- Ability to export materials containing information about holdings;
- Viewing statistics and progress of portfolio performance;
- Availability of translation of text comments into specified natural languages.

Functional testing was performed, which showed the completeness and correctness of the functionality of the system.

4 FEASIBILITY STUDY FOR THE DEVELOPMENT OF AN INTELLIGENT SYSTEM OF FORMING TEXT DESCRIPTIONS OF INVESTMENT ACTIVITIES

4.1 Characteristics of the system for forming text descriptions

The subject of the thesis project is the intellectual system of formation of text descriptions on investment activity. This system is designed for investment control with the possibility of automated generation of text comments.

The development and implementation of this application will:

- Conduct performance monitoring of holdings;
- Analyze the profitability of portfolios;
- Select an investment strategy suitable for the client;
- Compose textual descriptions based on available data to identify strengths and weaknesses of instruments.

The system is developed by individual order of the community of investment agents. The developers are employees of the closed joint-stock company "Qulix Systems".

4.2 Calculation of investments in intelligent system development

Calculation of software development costs in this case will consist of the following items:

- costs for the developers' main salaries; of developers' extra payroll;
- social security contributions;
- other costs (depreciation, electricity, travel expenses, rent for office space and equipment, management and sales expenses, etc.).

1. Costs of basic salaries of the development team. Calculation of basic salary costs of the development team is based on the composition and size of the team, the amount of monthly salary of each team member, as well as the labor intensity of work performed during the development of the system by individual performers and is carried out according to the formula:

$$Z_o = K_{\Pi\P} \cdot \sum_{i=1}^n Z_{\text{qi}} \cdot t_i, \quad (4.1)$$

where n – is the number of performers engaged in the development of specific software;

$K_{\Pi\P}$ – bonus ratio (1,5);

$З_{qi}$ – hourly wage of the i-th performer (rubles);

t_i – labor intensity of the work performed by the i-th performer (h).

Bonus ratio $K_{пп}$ is assumed to be 1.1. The hourly wage for each performer is determined by dividing his monthly wage by the number of working hours in the month, i.e., by 168 hours.

The basic wage calculation is presented in table 4.1. The data on wages for the development team is provided by Qulix Systems as of May 09, 2022.

Table 4.1 – Calculation of basic wages

№	Team member	Type of work performed	Monthly salary, rubles.	Hourly wage, rubles.	Labor-intensity of work, h.	Salary according to the rate, rub.
1	2	3	4	5	6	7
1	Project Manager	Analytics	2821,5	16,79	70	1175,63
2	Machine Learning Engineer	Development	2700	16,07	45	723,214
3	Tester	Testing	1890	11,25	120	1350
4	Front-end developer	development	2565	15,26	235	3587,95
5	Back-end developer	development	4050	24,10	260	6267,86
Total						13104,6
Bonus(50%)						1310,46
Total costs for the developers' basic salaries						14415,1

2. The cost of additional wages for the development team includes payments provided by labor legislation (payment for work leave, reduced hours, time for performing public duties and other payments not related to the main activity of the performers), and is determined by the formula:

$$З_{д} = \frac{З_{о} \cdot H_{д}}{100}, \quad (4.2)$$

где $H_{д}$ – additional wage norm (20 %);

$З_{о}$ – costs of basic wages and salaries, (rubles);

The additional wages will be:

$$З_{д} = \frac{14415,11 \cdot 15}{100} = 2162,27 \text{ руб.}$$

3. Contributions to social security and compulsory insurance (social security fund and compulsory insurance) are determined in accordance with current legislation according to the formula:

$$P_{\text{coи}} = \frac{(3_o + 3_d) \cdot H_{\text{coи}}}{100}, \quad (4.3)$$

where $H_{\text{coи}}$ – The rate of deductions to the Social Security Fund and Belgosstrakh (in accordance with current legislation – 34,6 %).

$$P_{\text{coи}} = \frac{(14415,11 + 2162,27) \cdot 34,6}{100} = 5735,77 \text{ руб.}$$

4. Other costs, such as employee training, are included in the cost of software development as a percentage of the development team's basic salary costs, calculated according to the formula:

$$P_{\text{np}} = \frac{3_o \cdot H_{\text{np}}}{100}, \quad (4.4)$$

где H_{np} – other expenses rate (30 %).

$$P_{\text{np}} = \frac{14415,11 \cdot 30}{100} = 4324,53 \text{ руб.}$$

5. The total development costs are calculated according to the formula:

$$3_p = 3_o + 3_d + P_{\text{coи}} + P_{\text{np}}, \quad (4.5)$$

$$3_p = 14415,11 + 2162,27 + 5735,77 + 4324,53 = 26637,67 \text{ руб.}$$

6. The expected profit, included in the price of the software, is calculated by the formula:

$$\Pi_{\text{nc}} = \frac{3_p \cdot P_{\text{nc}}}{100}, \quad (4.6)$$

where P_{nc} – is the cost-effectiveness of software development, equal to 25 %.

$$\Pi_{\text{nc}} = \frac{26637,67 \cdot 25}{100} = 6659,41 \text{ руб.}$$

7. The selling price of the software is calculated by the formula:

$$\Pi_{\text{nc}} = 3_p + \Pi_{\text{nc}}, \quad (4.7)$$

$$\Pi_{\text{nc}} = 26637,67 + 6659,41 = 33297,09 \text{ руб.}$$

Formation of the price on the basis of costs is shown in the table 4.2.

Table 4.2 – Formation of the price of software based on costs

Expenditure item	Sum, rubles
Basic salary of the development team	14415,11
Additional wages for the development team	2162,27
Contributions to social security and compulsory insurance	5735,77
Other expenses	4324,53
Total development costs	26637,67

4.3 The economic effect of the development and use of the intelligent system

Any software is developed to satisfy certain needs, to produce an effect [37]. In general, an effect can be of 2 kinds:

a) **Economic effect.** The development and use of software directly affects the economic performance of the user (for example, as a result of development the developer receives an increase in profit from the sale of software, automated quality control system significantly reduces losses from defects, thereby reducing the cost of production, and hence increasing profits). This effect is easy to cost estimate and must necessarily be calculated in the economic justification;

б) **Non-economic effect.** This is an effect not directly related to the economic results of the company's activity: social, environmental, political, technical. In this case, the use of software has an indirect (indirect) impact on the economic performance of the user, or provides additional benefits of another nature, which often cannot be evaluated in monetary terms, or the evaluation process is complicated, time-consuming and inaccurate.

The result in the use of the software product is an increase in net profit.

For the organization-developer, the economic effect is the increase in net profit received from the development and sale of the software product to the customer. The portal is sold by the organization-developer at a selling price formed on the basis of the development costs presented in the table 4.2. The economic effect obtained by the organization-developer in the form of an increase in net profit from its development is determined by the formula;

$$\Delta\Pi_{\text{nc}} = \Pi_{\text{nc}} \cdot \left(1 - \frac{H_{\Pi}}{100}\right), \quad (4.8)$$

where Π_{nc} – profit included in the price of the software tool.

Since the developer CJSC "Qulix Systems" is a resident of the High-Tech Park, it is exempt from value added tax and income tax ($H_{\pi} = 0$). The increase in net profit will be:

$$\Delta\Pi_q = 6659,41 \cdot \left(1 - \frac{0}{100}\right) = 6659,41 \text{ руб.}$$

4.4 Calculation of indicators of economic efficiency of the development and use of the intellectual system

For the organization-developer of software development cost-effectiveness assessment is carried out by calculating the simple rate of return (cost-effectiveness of software development costs) by the formula:

$$P_{\pi} = \frac{\Delta\Pi_{\pi c}}{Z_p} \cdot 100\%, \quad (4.9)$$

where $\Delta\Pi_q$ is the increase in net profit received from the development of the knowledge portal by the organization-developer by individual order;

Z_p – the cost of developing the system by the organization-developer.

Thus, the economic efficiency will be:

$$P_{\pi} = \frac{6659,41}{26637,67} \cdot 100\% = 25\%.$$

4.5 Conclusion

As a result of the feasibility study of the development of an intelligent system of forming text descriptions on investment activities the following values of performance indicators were obtained:

a) According to the results of calculations, the cost of developing the intellectual system is 26637.67 rubles.

b Increase in net profit is equal to 6659.41 rubles. Return on investment is 25 %.

Thus, the development and application of the software product is effective and these investments are worthwhile.

CONCLUSION

As a result of work on the diploma project the software solution in the form of intellectual system of formation of text descriptions on investment activity was developed.

The analysis of approaches to the development of systems designed to form financial comments, input and output information, characteristics and requirements to the system, analogues on the market.

In the process of designing the main entities and main types of users of the given system were singled out, the conceptual model of data was designed, main functional capabilities of the system and functions of users were described. The method for the implementation of automated text generation was selected and the stage of designing API for communication with language models was described.

In the process of development of the system the analysis of existing technologies used in the development of similar systems was conducted. The basic requirements to the used technologies were determined. Necessary functionality of the system and user interface, which provides the user with interaction with the developed system, were implemented.

The following features are implemented in the presented project:

- user management;
- portfolio and holding management;
- generation of reports and comments;
- ability to export materials containing information about holdings;
- viewing of statistics and progress of portfolio performance;
- availability of translation of text comments into specified natural languages.

Software implementation of the system is based on multilevel architecture using stack of modern client and server application development technologies: Python, Java, JavaScript, TypeScript, React, FastAPI, Spring. Presence in the system of API for communication with pre-trained models allows using the best methods of natural languages processing by using Transformers T0 and GPT-3.

The feasibility study of the development of the system has been made, which allowed to conclude the economic feasibility.

The application is planned to be used in the field of finance.

The system is a fully automated market commentary analysis tool for individual portfolios, based on a robust and efficient approach. All comments for individual portfolios are selected based on a modular approach according to performance and contribution, investment strategy and macroeconomic background. Asset managers can use it as a support or basis for further analysis of individual comments. The text modules can be automatically adjusted in length or content and adjusted to the client's level of financial knowledge.

In the future, it is planned to personalize the texts generated by the models, to automate the retraining of language models on the data corrected during the writing of reports.

The diploma project tasks are solved and the goals are achieved. Diploma project is completed and designed in accordance with the current enterprise standard and state standards [38].

REFERENCE LIST

- [1] Los, Cornelis A. Computational Finance: A Scientific Perspective / Cornelis A. Los. — World Scientific Publishing Company, 2000. — 340 P.
- [2] Alexander, G.J. Fundamentals of Investments / G.J. Alexander, W.F. Sharpe, J.V. Bailey. — 3rd edition. — Prentice Hall, 2001. — 781 P.
- [3] Guerard, J.B. Handbook Of Applied Investment Research / J.B. Guerard, W.T. Ziemba. — World Scientific Publishing Company, 2020. — 816 P.
- [4] Ashby Monk, Rajiv Sharma. Reframing Finance: New Models of Long-Term Investment Management / Rajiv Sharma Ashby Monk, Duncan L. Sinclair. — Stanford University Press, 2017. — 216 P.
- [5] Marcus Invest Review 2022 – Forbes Advisor [Electronic resource]. — Mode of access: <https://www.forbes.com/advisor/investing/marcus-invest-review/>. — Date of access: 06.03.2022.
- [6] Investing | Marcus by Goldman Sachs® [Electronic resource]. — Mode of access: <https://www.marcus.com/us/en/resources/investing>. — Date of access: 21.03.2022.
- [7] Financial Planning, Management, and Advice | Vanguard [Electronic resource]. — Mode of access: <https://investor.vanguard.com/advice/financial-advisor/personal-advisor-services>. — Date of access: 13.03.2022.
- [8] Vanguard Personal Advisor Services Review 2022 - Forbes [Electronic resource]. — Mode of access: <https://www.forbes.com/advisor/investing/vanguard-personal-advisor-services-review/>. — Date of access: 07.03.2022.
- [9] Bloomberg Terminal | Bloomberg Professional Services [Electronic resource]. — Mode of access: <https://www.bloomberg.com/professional/solution/bloomberg-terminal/>. — Date of access: 27.04.2022.
- [10] Bloomberg - Forbes [Electronic resource]. — Mode of access: <https://www.forbes.com/companies/bloomberg/>. — Date of access: 09.03.2022.
- [11] Investment Analysis Platform | Morningstar [Electronic resource]. — Mode of access: <https://www.morningstar.com/products/direct/emea>. — Date of access: 04.04.2022.
- [12] Speed Matters: How Morningstar Accelerates Innovation With Data [Electronic resource]. — Mode of access: <https://www.forbes.com/sites/amazonwebservices/2021/09/28/>

speed-matters-how-morningstar-accelerates-innovation-with-data/. — Date of access: 11.03.2022.

[13] JPMorgan Chase (JPM) - Forbes [Electronic resource]. — Mode of access: <https://www.forbes.com/companies/jpmorgan-chase/>. — Date of access: 15.03.2022.

[14] Wealth Management - JP Morgan [Electronic resource]. — Mode of access: <https://www.jpmorgan.com/wealth-management>. — Date of access: 01.04.2022.

[15] A survey on web services composition [Electronic resource]. — Mode of access: <https://doi.org/10.1504%2FIJWGS.2005.007545>. — Date of access: 24.03.2022.

[16] Orfali, Robert. Client/Server Programming with OS/2 2.1 (3rd Ed.) / Robert Orfali, Dan Harkey. — 3rd edition. — Van Nostrand Reinhold Co., 1993. — 1142 P.

[17] Unsupervised Word Sense Disambiguation Rivaling Supervised Methods [Electronic resource]. — Mode of access: <https://aclanthology.org/P95-1026>. — Date of access: 10.04.2022.

[18] Natural Language Processing and Computational Linguistics. Computational Linguistics [Electronic resource]. — Mode of access: https://doi.org/10.1162/coli_a_00420. — Date of access: 06.04.2022.

[19] The Illustrated Word2vec [Electronic resource]. — Mode of access: <https://jalamar.github.io/illustrated-word2vec/>. — Date of access: 11.04.2022.

[20] Efficient Estimation of Word Representations in Vector Space [Electronic resource]. — Mode of access: <https://arxiv.org/pdf/1301.3781.pdf>. — Date of access: 25.03.2022.

[21] Machine Learning Meets Natural Language Processing [Electronic resource]. — Mode of access: https://www.researchgate.net/publication/351046768_Machine_Learning_Meets_Natural_Language_Processing_-_The_story_so_far. — Date of access: 17.04.2022.

[22] What is a REST API? - Red Hat [Electronic resource]. — Mode of access: <https://www.redhat.com/en/topics/api/what-is-a-rest-api>. — Date of access: 11.04.2022.

[23] DeepL API [Electronic resource]. — Mode of access: <https://www.deepl.com/docs-api/>. — Date of access: 02.05.2022.

[24] The Tale of T0 - BigScience [Electronic resource]. — Mode of access: <https://bigscience.huggingface.co/blog/t0>. — Date of access: 17.04.2022.

[25] OpenAI API [Electronic resource]. — Mode of access: <https://beta.openai.com/docs/introduction>. — Date of access: 14.04.2022.

[26] Wiederhold, Gio. A structural model for database systems / Gio Wiederhold, Ramez A. Elmasri. — 1st edition. — Stanford University, 1979. — 51 P.

[27] TypeScript: JavaScript With Syntax For Types. [Electronic resource]. — Mode of access: <https://www.typescriptlang.org>. — Date of access: 01.05.2022.

[28] React - A JavaScript library for building user interfaces. [Electronic resource]. — Mode of access: <https://reactjs.org/>. — Date of access: 01.05.2022.

[29] The OpenAPI Specification Repository - GitHub [Electronic resource]. — Mode of access: <https://github.com/OAI/OpenAPI-Specification>. — Date of access: 14.03.2022.

[30] Spring Boot [Electronic resource]. — Mode of access: <https://spring.io>. — Date of access: 29.04.2022.

[31] Welcome to Python.org [Electronic resource]. — Mode of access: <https://www.python.org/>. — Date of access: 13.03.2022.

[32] FastAPI [Electronic resource]. — Mode of access: <https://fastapi.tiangolo.com/>. — Date of access: 09.04.2022.

[33] Multitask Prompted Training Enables Zero-Shot Task Generalization [Electronic resource]. — Mode of access: <https://arxiv.org/pdf/2110.08207.pdf>. — Date of access: 05.04.2022.

[34] Language Models are Few-Shot Learners [Electronic resource]. — Mode of access: <https://arxiv.org/pdf/2005.14165.pdf>. — Date of access: 19.03.2022.

[35] Silberschatz, Abraham. Database Systems Concepts / Abraham Silberschatz, Henry F. Korth, S. Sudarshan. — 4th edition. — McGraw-Hill Higher Education, 2001. — 1088 P.

[36] PostgreSQL: The world's most advanced open source database [Electronic resource]. — Mode of access: <https://www.postgresql.org/>. — Date of access: 07.03.2022.

[37] Горовой, В. Г. Экономическое обоснование проекта по разработке программного обеспечения / В. Г. Горовой, А. В. Грицай, В. А. Пархименко. — Минск : БГУИР, 2018. — 12 с.

[38] Доманов, А. Т. СТП 01-2017. Стандарт предприятия. Дипломные проекты (работы). Общие требования / А. Т. Доманов, Н. И. Сорока. — Минск : БГУИР, 2017. — 169 с.