Automation of Visualization of Indicators of Product Quality Assessment at an Industrial Enterprise

Rauza R. Abdulveleeva
Mathematics and Science dept.
Novotroitsky branch of the
"National Research Technological University MISIS"
Novotroisk, Russia
rashitovna-2011@mail.ru

Ildar R. Abdulveleev
Power supply of industrial enterprpises dept.
Nosov Mangnitogorsk State Technical University
Magnitogorsk, Russia
i.abdulveleev@magtu.ru

***Abstract*—The article considers the problem of developing a solution for automating the process of visualizing indicators of the quality assessment of manufactured products at an industrial enterprise. The existing methods and tools for data visualization are considered, the architecture and main components of the proposed automation system are presented, and the advantages of the proposed solution compared to traditional approaches are discussed. A description of the process of creating a computer program designed to automate the processing of big data and their visualization is proposed. The structured data entering the program are metal quality indicators, which are obtained on the basis of processing reporting documentation from the enterprise information system. The features of software libraries and methods of working with them are described, and an algorithm for the software module is developed. Data visualization in the form of graphs is based on information about faulty products, presented as multidimensional data.**

Keywords—visualization, dashboard, metal defects, graphic visualization, web application

I. INTRODUCTION

Ensuring high quality of manufactured products is one of the key tasks for any industrial enterprise. Quality control at all stages of the production process allows identifying and eliminating potential defects, reducing the level of defects and increasing customer satisfaction. However, effective quality management is impossible without a clear understanding and analysis of the relevant indicators characterizing various aspects of production. Traditionally, product quality assessment at enterprises was carried out by collecting and processing data manually, which required significant time and labor costs. In addition, presenting the results in the form of tables and simple graphs often made it difficult to identify hidden patterns and trends necessary for making informed management decisions.

In modern conditions, when the volume of production and the number of controlled parameters are constantly growing, there is an urgent need to automate the processes of collecting, processing and visualizing data on product quality. Visualization of quality assessment indicators using modern tools and methods allows you to present complex data sets in a clear and easily interpretable form, facilitating analysis and identifying critical areas requiring intervention. Traditional data visualization methods such as tables, line graphs, histograms and pie charts are widely used to present product quality indicators. These methods are easy to create and understand, but have limited capabilities for displaying large and multidimensional data sets [1].

With the advancement of technology and the growth of data volumes, more advanced visualization tools have emerged:

- interactive dashboards: Tools such as Tableau, Power BI, and Qlik allow you to create interactive dashboards with different types of visualizations, filters, and drill-down capabilities;
- business intelligence (BI) tools: BI solutions such as Microsoft BI, SAP BusinessObjects, and IBM Cognos include powerful data visualization and analytics capabilities;
- specialized visualization software: programs such as Visme, Infogram, and Piktochart are designed specifically for creating visual representations of data, infographics, and presentations
- open source libraries and frameworks: JavaScript libraries such as D3.js, Plotly.js, and Chart.js provide flexible tools for creating interactive data visualizations on web pages.

Modern data visualization tools offer a number of advantages, including increased visibility, interactivity, the ability to perform deep analysis, and identify hidden patterns. However, they can also be difficult to learn and implement, and require significant licensing and support costs [2, 3]. When choosing a suitable visualization tool, it is necessary to consider criteria such as functionality, scalability, compatibility with existing systems, ease of use, and cost of ownership.

Data visualization based on the use of Big Data provides powerful tools for their visual presentation, which helps to better understand correlations, data visualization using graphs and charts allows you to quickly identify relationships. The use of interactive dashboards allows users to see correlations and trends in real time, which facilitates decision making.

Visual representation of quality metrics enables quick identification of problem areas, trends, and patterns that may be difficult to detect when analyzing raw data or tables. According to research conducted by Edward Tufte, one of the leading experts in the field of data visualization, “properly

979-8-3315-1124-1/25/$31.00 ©2025 IEEE

created visualization is a tool for reasoning about evidence and data" [4]. Visualization helps overcome the limitations of human perception and information processing, allowing for quick and intuitive understanding of complex data sets.

In the context of product quality control in industrial plants, visualization of quality metrics can provide the following benefits:

- Quick identification of problem areas and anomalies. Visual representation of data allows easy identification of areas with high levels of defects, waste, or other undesirable phenomena [5].
- Analysis of trends and patterns. With the help of visualization, it is possible to track changes in quality metrics over time, identify seasonal or cyclical fluctuations, and determine correlations between different factors [6].
- Improved decision-making processes. Visual representation of data facilitates understanding of the situation and enables managers and quality specialists to make informed decisions on corrective and preventive actions [7].
- Improved communication efficiency. Data visualization is a universal language that enables effective communication of product quality information to various stakeholders, including management, production personnel, and customers [8].

Data visualization is therefore a key tool for effectively analyzing and interpreting large amounts of information, especially in critical areas such as product quality control in industrial enterprises. Visualization of quality assessment metrics in industrial enterprises allows for more efficient quality control, timely detection of problems, and informed decision-making aimed at improving production processes and increasing customer satisfaction.

In recent years, the Big Data concept has become a key tool in various industries, including metallurgy. Big Data allows you to collect, store and analyze huge amounts of data, which contributes to a deeper understanding of processes, including the study of product defects. Even with significant success in digital transformation, industrial enterprises still do not sufficiently use the full power of Big Data. This depends on several factors: first of all, the need to digitize data from equipment and machines (in other industries, as a rule, this factor simply does not exist). The number of data sources is growing rapidly, which means that the development of technology for their processing is becoming increasingly in demand.

Implementation of Big Data requires both technical and organizational support. The former involves organizing data extraction, data storage, unified automated workplace analysis, creating digital models, optimization and forecasting. Metallurgical enterprises generate large volumes of data at each stage of the production process. These data may include: production data (temperature, pressure, rolling speed and other parameters); quality data (results of strength tests, corrosion resistance and other characteristics); defect data: information on the types, quantity and causes of defects.

Using Big Data technologies allows you to integrate and analyze this data in real time, which helps identify patterns and trends. Big Data also helps optimize production processes. For example, by analyzing defect data, companies can identify bottlenecks in the production chain and make changes that will help reduce the level of defects. This may include adjusting equipment, changing process parameters, or improving quality control.

Product quality in metallurgy is critically important, as it directly affects the safety, durability, and efficiency of metals in various industries. However, in metallurgical enterprises, defects inevitably occur during production, that is, individual product non-conformities with regulatory requirements. All defects in connections are of two types: external or internal. In modern production, there are plans for testing the state of defects. Slices are taken according to the plan on a regular basis and this is always associated with big data.

According to Rosstat data and industry reports, the number of defects in metallurgical products can vary depending on the type of production and technology. For example, in 2021 reports, it was reported that the level of defects at Russian metallurgical enterprises ranged from 5% to 15% depending on the specific product and stage of production (e.g. casting, rolling).

Various types of defects can occur at metallurgical enterprises, including: casting defects (include cracks, porosity, inclusions and other defects that arise during the casting process); rolling defects (such as unevenness, deformations, cracks and others that can occur during metal processing); corrosion defects (occur as a result of exposure of the metal to the external environment, which can lead to the destruction of products).

According to a study conducted in 2020 by the Institute of Metallurgy and Materials Science, the number of identified defects in metallurgical products in Russia has increased by 10% over the past five years. In particular, in 2019, more than 1,200 cases of defects were recorded at large metallurgical enterprises such as NLMK and Mechel.

According to Rosstat, in 2021, the level of defects in the metallurgical industry amounted to about 10% of the total volume of manufactured products. This means that out of 10 million tons of metal, about 1 million tons were identified with various defects. An analysis of defects at metallurgical enterprises in Russia shows that the level of defects in the industry ranges from 8% to 13%. The main causes of defects are related to the quality of raw materials, technological errors and insufficient qualifications of personnel. To reduce the level of defects, it is necessary to introduce modern technologies, improve quality control and improve the qualifications of workers.

An analysis of the number of defects at metallurgical enterprises in Russia shows that the problem of product quality requires constant attention and improvement. The introduction of modern technologies, personnel training and improvement of quality control can significantly reduce the level of defects.

II. Description of the Application

To find out the reason for the systematic occurrence of related types of defects, a large number of measuring devices are installed on the equipment. Initially, various types of information come from the sensors to the enterprise's own server. The next step is to examine the received data for signs of systematicity and analyze it. Depending on the enterprise, this happens in automatic or semi-automatic modes.

The result of the defect analysis is a large array of data, the information in which is unstructured, that is, presented in an unclear and incomprehensible form. Most enterprises generate reporting documentation manually or semi-automated. The entire quality management department works on compiling a report on defects based on the analysis results, and this often takes a lot of time. Then the received information is structured, and employees create dashboards from it - interactive reports, visual monitoring. This process is time-consuming and labor-intensive in terms of the amount of information processed.

The specialist who compiles this documentation sometimes needs non-trivial knowledge of the described area. After the documentation is generated, the compiled report is checked many times and, despite this, may contain errors. The developed software product is designed to automate the process of report preparation and reduce human participation to a minimum. After new information is received in the database, the already generated report is automatically changed and will be available online from the enterprise portal.

The developed application for automating the process of visualizing data obtained from the results of assessing the quality of manufactured products of an industrial enterprise is implemented using Big Data technology. The application automates the process of structuring information and creates a dashboard. Updated information on defects is added to the general database, due to which the dashboard is updated in real time.

The structured data are the quality indicators of metal products, which are obtained based on the processing of reporting documentation in the enterprise information system. One of the tools for automating the process of visualizing the quality indicators of metals and defects in the metallurgical industry is the use of open libraries for visualization and data analysis:

- The Python Matplotlib library is widely used for data visualization in metallurgy, including the construction of graphs, histograms and control charts [9].
- The R ggplot2 library provides powerful tools for creating informative graphs and visualizations for analyzing the quality of metals [10].

In case of large arrays of data on the quality of the manufactured products of the enterprise, graphical visualization is simply necessary.

A method of data presentation in the form of graphs in the Python environment is proposed. The following libraries were selected for implementation: Matplotlib, Pandas, Numpy, Streamlit. Using packages for working with Python libraries helps to diversify the visualization of the data under consideration and, most importantly, to make graphical objects more informative.

Matplotlib is a popular library for data visualization written in Python. The library has full customization of data, parameters, graphs and rendering. Matplotlib consists of many modules. The modules are filled with various classes and functions that are hierarchically related to each other. During the development of the software, the matplotlib.pyplot function collection was used.

The matplotlib.pyplot interface is a set of commands and functions that make the syntax of graphical matplotlib commands similar to the commands used in the MATLAB environment, where one environment has tools for both drawing and numerical analysis. Matplotlib has pylab, which combines the pyplot and numpy modules into a single namespace.

Before starting development, a set of data should be formed to which the program logic will refer. The following types of information flows are accepted:

- input information flows: shipment date, specific gravity, organization name, product type, product purpose, receipt date, declared weight, defect name, defect group, recognition date, recognized weight, receipt type, declared amount, recognized;
- output information flows by product type: share of declared from shipped, by shipment date; share of declared from shipped, by receipt date; share of recognized from declared, by receipt date, in rubles; share of recognized from declared, by receipt date, in tons; share of recognized from shipped, by shipment date; share of recognized from shipped, by recognition date; declared tonnage, by shipment date; declared tonnage, by receipt date; tonnage recognized, by date of shipment; tonnage recognized, by date of recognition;
- output information flows by combinations: total; defect; product type; product type and defect; defect group; defect group and product type; purpose; purpose and defect; purpose and product type; purpose, product type and defect; purpose and defect group; purpose, product type and defect group.

There are no special requirements for the characteristics of the information system. The input data are the readings obtained in the daily reports on defective products. The report is sent to the enterprise information system and to the database, including a comparison of the amount of spoiled products to date with the amount of spoiled products in previous days. The brands of defective products, workshop, time and cause of the malfunction are also indicated.

Since in the designed database (Figure 1) the relationship between two objects occurs on a one-to-one basis, only one table is used. This reduces the number of connections that need to be performed in the database to obtain results. The parameters of the developed database are the input information flows described above. The reports proposed by the quality management department of one of the Russian ferrous metallurgy enterprises were used as initial data.

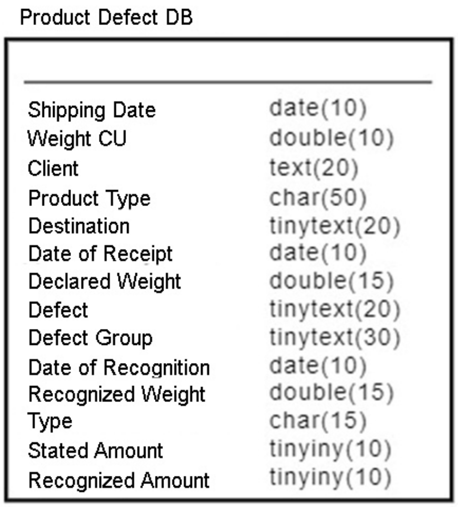

Fig. 1. Database structure.

One of the goals of software development is to quickly and automatically generate a report on product defects. To achieve this goal, it is necessary to reduce the number of parameters as much as possible. As an example, to test the performance of the program, a database consisting of one hundred records and fourteen fields is considered. During the program's operation, more than 1.5 thousand different graphs are formed in just a few seconds, whereas with a traditional approach at an industrial enterprise, such data visualization would take about eight hours.

The application uses the original Oracle Database, which is an object-relational system. It is supported by technologies that implement an object-oriented approach, provides management of the processes of creating and using databases. During the process of creating a database, it is necessary to prepare a number of operating system files and use them together as a single database. The database is created once, regardless of the number of files in which it is located, and does not depend on how many requests to it will be taken into account. All this data must be systematized and monitored using a Python script. In case of frequent malfunctions for the same reason, a report should be sent to the relevant department. Figure 2 shows the program's operation algorithm. The dashboard's visual interface allows users to receive, filter and analyze data provided by various workshops of the enterprise. The proposed system allows data exploration using a simple query interface and a set of flexible interactive visualization tools.

III. Visualization Application Operation

Visualization is central to monitoring and creating dashboards. Incoming input data is transformed, if necessary, using the Pandas and Numpy libraries. Pandas is used to format data, for example, in converting date and time formats. Numpy is used for mathematical functions or casting data arrays as axes.

The web application is deployed on a local server. To solve this problem, the Streamlit framework was used. Streamlit is a web platform for creating interactive data visualization panels. In essence, dashboards. The framework is optimized for working with data and supports the Pandas, Numpy, and Matplotlib libraries. The main advantage of this framework is that it does not require additional intervention in the program body. That is, optimization for data libraries separates the development of a web application from working with data. And this will allow you to quickly switch to another framework if necessary [11-15]. A fragment of the program code is shown in the Figure 3.

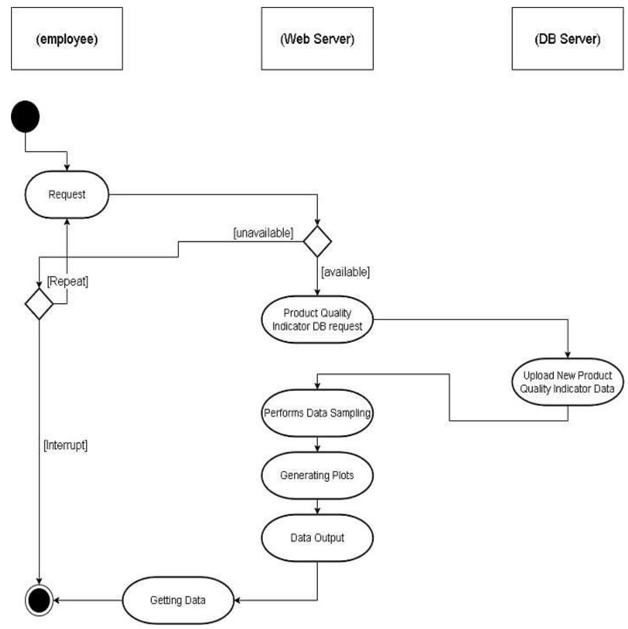

Fig. 2. Application operation algorithm.

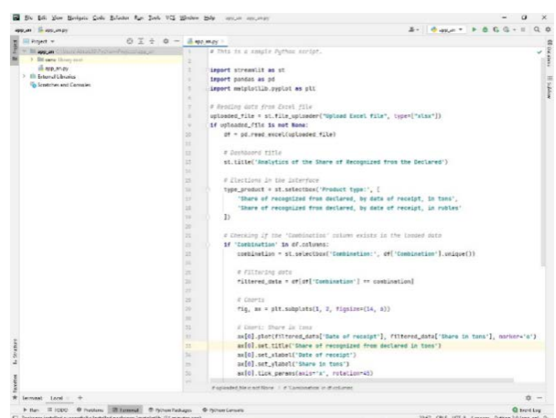

Fig. 3. A fragment of the working code of the application.

After launching the application, the user is prompted to go to the specified address in the “Terminal” field of the program console.After launching the application, the user is prompted to go to the specified address in the “Terminal” field of the program console (Figure 4). After activating the specified address, the application will open in the browser with a user interface created using Streamlit. A browser window will appear in the application (Figure 5), which will allow the user to upload an Excel file. The user selects a file to upload, which will then be processed by the program and will allow the visualization results to be displayed in the dashboard. For the visualization of graphs, for example, an original report in .xls format is required, containing data on defects. This report includes 14 features related to different types of products and is a time series.

Once loaded, the file will be read using the pandas library. The code in the application may contain logic for transforming or filtering data. For example, the program has code that allows you to select specific columns or rows. You can filter data in a DataFrame to selectively select individual rows or columns of interest to the user.

The loaded data is displayed on the dashboard using Streamlit functions. Streamlit is an open-source library designed for quickly developing interactive dashboards. It allows you to visualize data and create intuitive interfaces for analyzing it. With Streamlit, programs are developed that automatically generate visualizations and analytical graphs based on input data. This simplifies interpretation and analysis, since the indicators entering the system are adapted to the tasks and competence of the user.

The program provides data visualization using Matplotlib , which allows you to build graphs. Using the "plt.show()" function, you can display the graph directly on the dashboard. Working with data begins with reading it using pandas.read_excel. The resulting DataFrame is used for further processing, where basic interface elements are set, such as the title via streamlit.title and the drop-down menu using streamlit.selectbox. After that, the data is filtered according to the user's selection in the selectbox.

The IT department regularly receives reports with various indicators. The dashboard visualizes time series demonstrating the dynamics of the volumes of recognized tons of products by different parameters. The user can select parameters such as "shipment date" or "recognition date" and quickly analyze the received data. Working with the dashboard and Excel files in Streamlit provides powerful tools for interactive data analysis, which allows you to quickly develop and update visualizations by interacting with the loaded data.

Since real Excel files with reliable information about the manufactured products at the metallurgical enterprise are not publicly available, this information is closed and access to it is limited, to demonstrate the operability of the program, files were used in which the numerical data were generated randomly. Taking this fact into account, Figure 6 demonstrates the operation of the program, for example, in terms of visualization in tons of the recognized defect. Graphic display of information allows you to visualize the obtained data, perform an analysis, support the formulated conclusion, or emphasize an emphasis. The indicators of the share of the recognized defect from the declared, by the date of receipt in tons and in rubles are shown in Figure 7. Figure 8 shows a graph of the visualization of the indicators of the declared tonnage by the date of shipment and by the date of receipt.

Conclusions

During the work, the process of visualization of indicators for assessing the quality of manufactured products at one of the metallurgical enterprises was automated. The technological capabilities of the tools used in the program allow visualization and analysis of huge arrays of data and display them in the form of graphs. This development offers one of the solutions to the problem of automating the processing of the obtained results on the quality of manufactured metal products based on the use of Big Data technology. Displaying the results on the dashboard allows improving the information interaction of employees at the enterprise, processing production information, which comes in large quantities from various departments and from facilities.

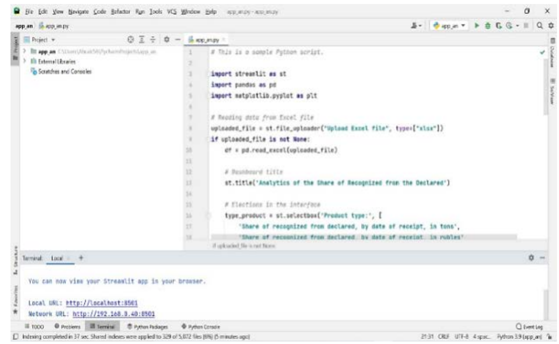

Fig. 4. Link for transition.

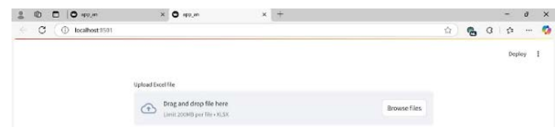

Fig. 5. Browser window.

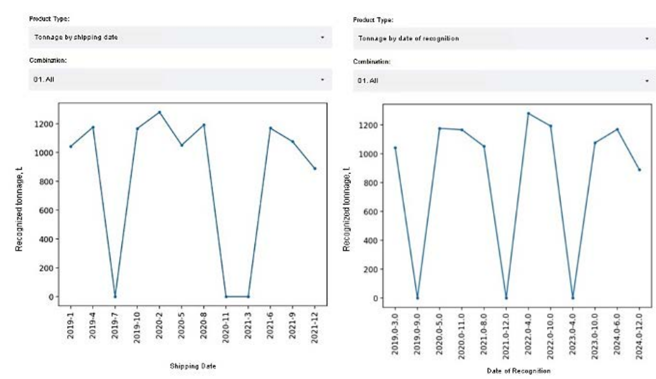

Fig. 6. Tonnage recognized by shipment date on the left chart and by recognition date on the right chart.

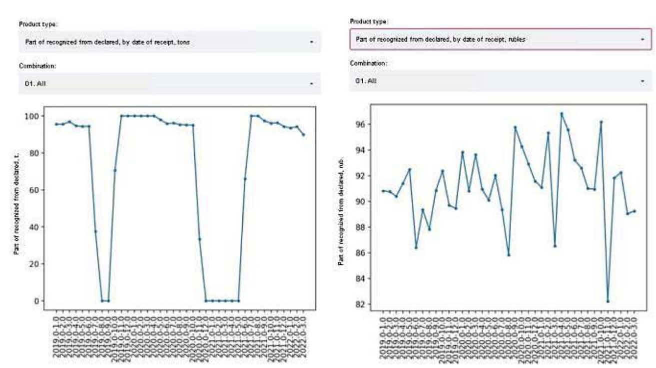

Fig. 7. Share of recognized from declared, by date of receipt in tons (left graph) and in rubles (right graph).

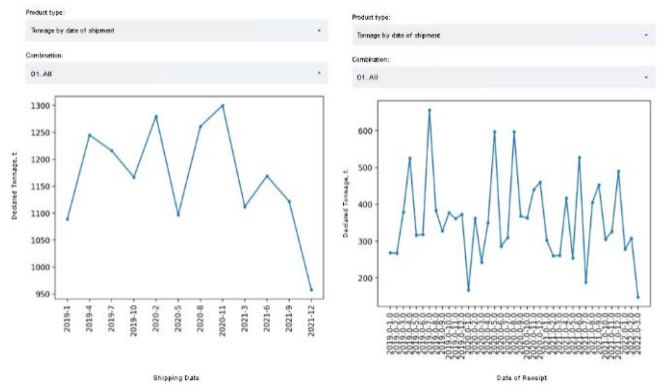

Fig. 8. Declared tonnage, by date of shipment (left graph) by date of receipt (right graph).

REFERENCES

[1] S. Few, Show Me the Numbers: Designing Tables and Graphs to Enlighten. Analytics Press, 2009.

[2] N. Zhu, Data Visualization with D3.js Cookbook. Packt Publishing, 2017.

[3] A. C. Telea, Data Visualization: Principles and Practice, CRC Press, 2014.

[4] E. R. Tufte, The Visual Display of Quantitative Information. Graphics Press, 2001.

[5] D. A. Keim, “Information visualization and visual data mining,” in IEEE Transactions on Visualization and Computer Graphics, vol. 8, no. 1, pp. 1–8, Jan.-March 2002, doi: 10.1109/2945.981847

[6] K. Healy, Data Visualization: A Practical Introduction. Publisher, 2018.

[7] R. R. Abdulveleeva, V. A. Korsakov, and I. R. Abdulveleev, “Crane movement modulator for web application of electric steel melting shop mnemonic scheme with crane movement visualisation system,” in 2024 International Ural Conference on Electrical Power Engineering (UralCon), 2024, pp. 812–816, doi: 10.1109/UralCon62137.2024.10718976.

[8] M. I. Abdrakhmanov, Python. Data Visualization. Matplotlib. Seaborn. Mayavi, 2020.

[9] H. Wickham, ggplot2: Elegant Graphics for Data Analysis. Springer, 2016.

[10] J. Zelazny, Speak the Language of Diagrams, St. Petersburg: Mann, Ivanov and Ferber, 2010.

[11] Mitchell Ryan, Scraping Websites with Python. Moscow: DMK Press, 2016.

[12] E. Matiz, Learning Python. Game Programming, Data Visualization, Web Applications. Piter, 2019.

[13] R. R. Abdulveleeva and E. M. Ivanin, “Software for the analysis of data on defects of metal products,” in Mathematical and Software Systems in Industrial and Social Spheres, vol. 9, no. 1, pp. 33–36, 2021, doi: 10.18503/2306-2053-2021-9-1-33-36.

[14] R. R. Abdulveleeva and E. M. Ivanin, “Using the Streamlit library for quick assembly and deployment of the web application as a data visualization and dashboard creator,” in Actual Problems of Modern Science, Technology and Education: Abstracts of the 80th International Scientific and Technical Conference, Magnitogorsk, April 2022, pp. 340.

[15] L.V. Radionova, A.D. Chernyshev, “Mathematical Description of AFE Rectifier Closed Loop System,” Procedia Engineering, vol. 129, pp. 16–21, 2015.