Decision Support System for scientific and technical expertise

Belov Aleksandr V.
Department of the Applied Mathematics
National Research University "Higher School of Economics"
Moscow, Russia
avbelov@hse.ru

Bikbaev Bulat I.
Department of the Computer Engineering
National Research University "Higher School of Economics"
Moscow, Russia
bibikbaev@miem.hse.ru

Gevondyan Margarita S.
Interdistrict Inspectorate of the Federal Tax Service of Russia for the largest taxpayers № 1
Moscow, Russia
m.gevondyan.r9977@tax.gov.ru

Levitan Darya A.
Department of the Applied Mathematics
National Research University "Higher School of Economics"
Moscow, Russia
dalevitan@edu.hse.ru

Panina Irina Yu.
Department of the Applied Mathematics
National Research University "Higher School of Economics"
Moscow, Russia
iyupanina@edu.hse.ru

***Abstract*— Currently in the Russian Federation the state actively supports IT companies that are engaged in research and development work (R&D) as part of the implementation of the state program to create a digital economy. One of the forms of state support for IT companies is the provision of tax benefits to them. In this regard a very important process of verifying the validity of the provision of tax benefits is the examination of scientific and technical documentation provided by IT companies to the tax authorities. The paper considers the issues of creating a support system for conducting scientific and technical examinations of materials submitted by IT companies to confirm the correctness of classifying the work performed as R&D.**

In this article, we offer the specialized software named Support System which allows taxpayers to provide information about completed R&D projects in a checklist format in a declarative manner. Criteria for compliance of implemented IT projects with R&D projects are defined. The reporting documentation undergoes a multi-stage check in accordance with the developed algorithms for loading and analyzing text materials for compliance with the established criteria. The adequacy of used algorithms is confirmed by the results of computer simulation. The developed prototype of the support system for conducting scientific and technical examinations is planned to be introduced into the activities of tax inspections.

***Keywords*— digital economy, R&D projects, taxpayer, decision support system, scientific report, text mining**

I. INTRODUCTION

Science, especially fundamental science, refers to activities that cannot fully focus on market relations, as evidenced by all international experience. It does not fall under commercial criteria, therefore, the state in all developed countries assumes responsibility for financing fundamental research and promising developments that can have a revolutionary impact and ensure competitiveness. With the development of computer technology in the world scientific IT projects have become in demand. In many countries mechanisms for providing various preferences and benefits are used for government support of developing sectors of the economy and high-tech companies. For example, in the USA, a number of government departments are responsible for supporting fundamental research in the country, including the US National Science Foundation (NSF), whose functioning is entirely devoted to the development of science and technology. Decisions on the allocation of funds are made on the basis of an independent examination carried out by scientists from all over the country [1]. Also there is, for example, S&T, which deals with scientific and technical expertise in the field of national security. In Sweden, which occupies one of the leading places in terms of the development of science and technology, the dominant position in the R&D of the private sector is occupied by investments of large multinational groups of manufacturing enterprises, the state accounts for a relatively large share of R&D. Public funding is linked to the progress of the work and depends on the final results of the project, assessed on the basis of pre-established indicators [2]. In addition, such expertise is carried out by the Federal Research Center for Project Evaluation and Consulting Services only in the field of science, technology and innovation. This shows the need for scientific and technical expertise not only in Russia for specific areas, such as tax and patent, but also in world practice.

Over the past 15 years the Russian Federation has been developing a system of state support measures for organizations operating in the field of IT to eliminate technological dependence on foreign developments and stimulate new independent developments to ensure the competitiveness of domestic products.

Significant measures of government support include tax benefits provided for by the Tax Code of the Russian Federation [3].

For example, paragraph 7 of Article 262 of the Code establishes the right of a taxpayer who spends on R&D according to List 1 to include these expenses in the other expenses of the reporting (tax) period in which such research or development (separate stages of work) were completed or the initial cost of amortized intangible assets in the amount of actual costs using a coefficient of 1.5.

Amendments to Article 262 of the Code, allowing the use of an increased coefficient when attributing R&D costs to corporate income tax expenses, were introduced in order

979-8-3503-9759-8/22/$31.00 ©2023 IEEE

to implement the innovative scenario of economic development by the President of the Russian Federation by Federal Law №158-FZ of 22.07.2008 [4].

In order to create a mechanism of administration of R&D costs by the tax authorities in accordance with the list set by the Government of the Russian Federation to be taken into account in the amount of actual costs with a coefficient of 1.5, it is envisaged that the taxpayers who use the mechanism specified in the bill will have to submit to the tax authority at their registration location a report on R&D performed in accordance with the requirements laid down in the national standard (Inter-State Standard GOST 7.32-2001) [5].

Based on the above the tax benefit with the use of a coefficient of 1.5, provided for in paragraph 7 of Article 262 of the Code, is aimed at stimulating taxpayers who carry out work that has scientific novelty for the development of national science and technology as a whole.

At the same time there are many cases when a taxpayer unlawfully applied the benefit provided by Article 262 of the Code, which is revealed by the results of tax audits.

Thus, the tax authorities check the reports provided by taxpayers on the R&D performed. At the same time, the check is carried out in two directions:

- financial documentation accompanying the taxpayer's implementation of projects declared as R&D. Checking financial documents allows tax officer to determine the correctness of the calculation of taxes taking into account the relevant benefits and preferences;

- scientific and technical reports in order to determine the signs of scientific novelty and innovativeness of R&D, compliance of the work performed with the important scientific and technical directions established by the government.

In order to verify the reports it became necessary to create a system that would automate the process of documentation verification. The system being developed is created for the benefit of tax authorities and taxpayers as they participate in the distribution of benefits. The availability of such a software product will significantly reduce labor costs of the tax authorities in the analysis of submitted documents submitted by taxpayers, increase the quality of the results of control and analytical work and determine the prospects of legislative regulation of this procedure.

II. Problem Statement

The designed system has the following list of technical requirements:

• Integration with the federal tax service system.

• Web design looks like a federal tax service website design and is adaptive to any device.

• Access to the personal account for taxpayers and tax inspectors.

• Ability to work with a checklist and files.

• Support the analysis of the report, which is provided by the taxpayer for the examination for the benefit for the purposes of R&D.

• Formation of the act of expertise.

For the above reasons to work with such reports, we need a system capable of implementing the primary analysis of reports in accordance with the set criteria in the check-list. Thanks to the implemented system and an intuitive interface on the developed WEB site, any specialist conducting such an examination can obtain the required analysis result.

In the process of using the site the registration and authorization windows will appear for the user. After that client will go to the page of applying for expertise.

After the end of the examination the user will see the result with the present and absent elements of the report, a diagram of the frequency analysis of words on the subject encountered in the submitted document, as well as the result of assigning the report to one of the topics of research.

The program allows users to download files with ".pdf" and ".docx" extensions. The results of the analysis are opened on the page of analysis and results.

The roles and base interaction in our system are presented at Fig.1.

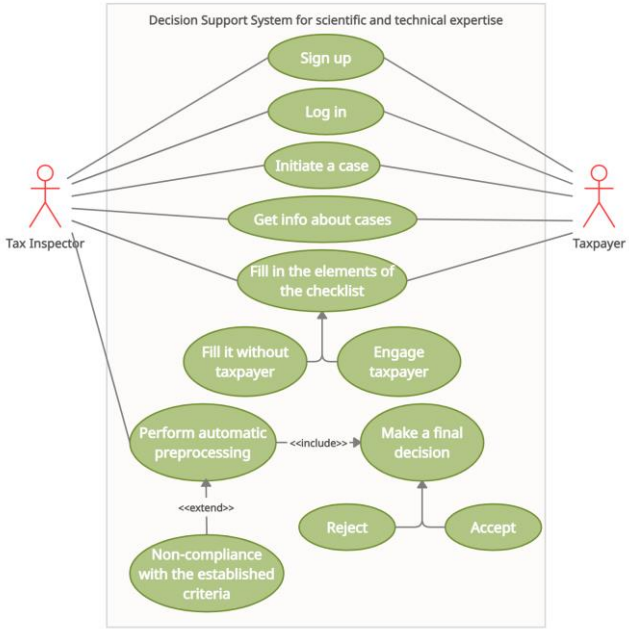

Fig. 1. UML Use Case Diagram

III. Implementation of the Developed System

The application is written in the python programming language using Django web framework [6]. This tool provides fast development, large amounts of documentation, secure and scalable. Django is based on the MTV pattern, which means Model, Template and View. Model interacts with the database and handles all the data operations, template processes how the data would be presented to the user, view describes what data will be provided upon request to a specific URL. UML class diagram for our models is presented at Fig.2.

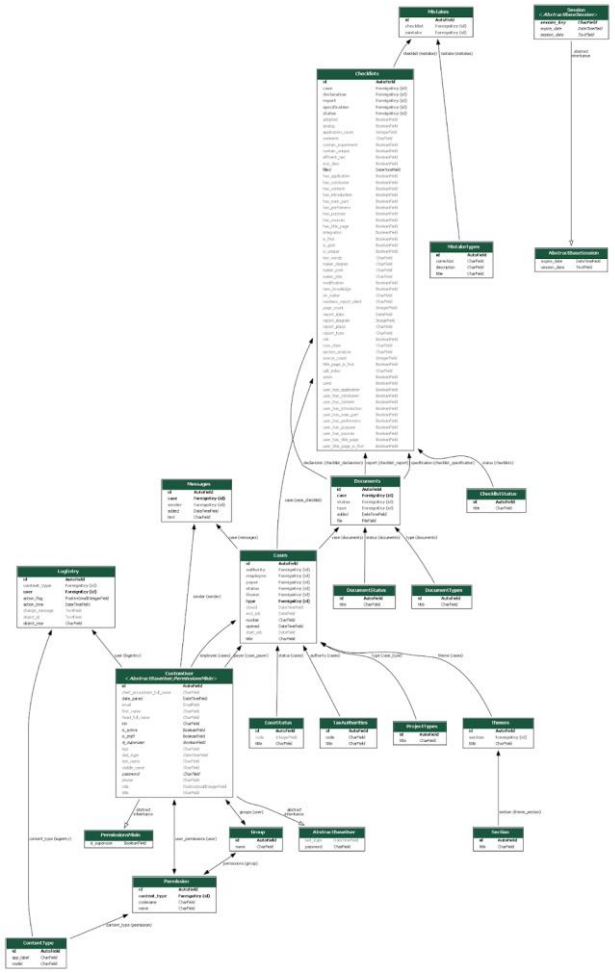

Fig. 2. UML class diagram for Database

The main technologies and their implementation in the system.

• Django as a framework.

• Bootstrap for developing a responsive and mobile-friendly website [7].

• PostgreSQL responsible for data storage. It is a powerful and reliable object-relational database system [8].

• GUnicorn as a WSGI server [9].

• NGINX is a scalable, secure, and reliable web server [10].

• The main programming language for textual analysis was Python, as well as its libraries for data analysis and visualization, such as nltk [11], seaborn [12], pandas [13], numpy [14], matplotlib [15].

System architecture is presented at Fig.3.

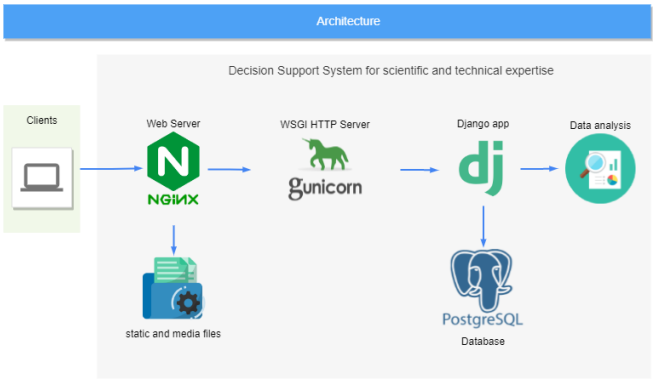

Fig. 3. System architecture

The stages of textual analysis are presented at Fig.4.

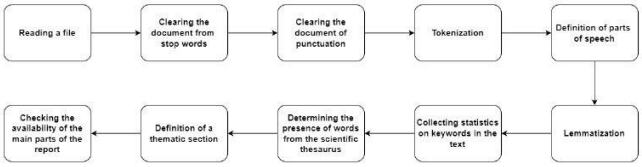

Fig. 4. The stages of textual analysis

Data analysis is located in a separate django view and called with the jQuery Ajax. It is an asynchronous method to send HTTP requests without waiting for a response.

User interface is presented at Fig.5 – Fig.10.

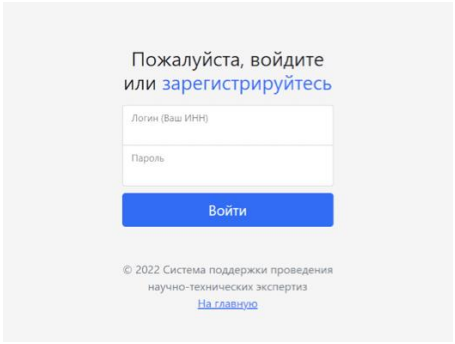

Fig. 5. User interface (screenshot 1)

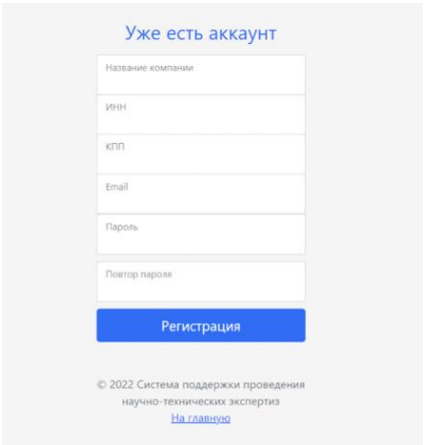

Fig. 6. User interface (screenshot 2)

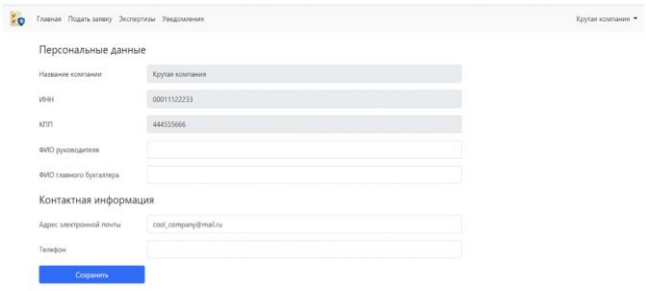

Fig. 7. User interface (screenshot 3)

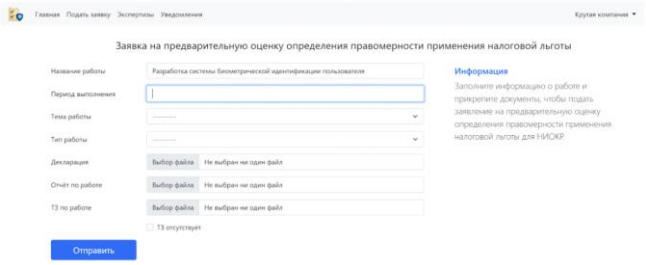

Fig. 8. User interface (screenshot 4)

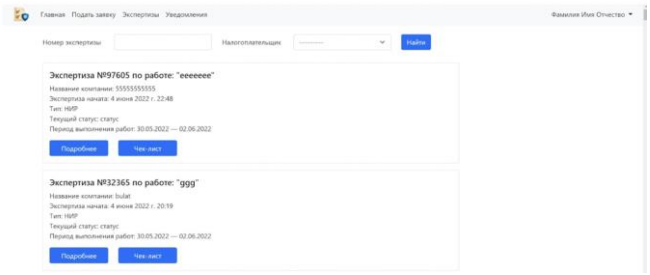

Fig. 9. User interface (screenshot 5)

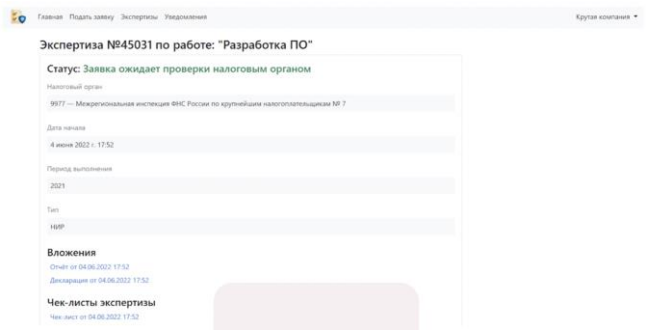

Fig. 10. User interface (screenshot 6)

Example of the dashboard, which shows information about the number of scientific words most commonly used in the report, is presented at Fig.11.

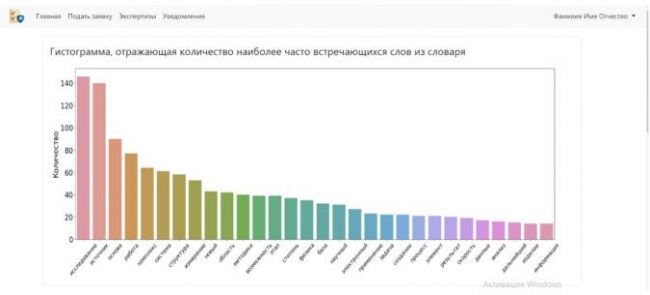

Fig. 11. Dashboard screenshot 1

Information about the structure of the report and assigning it to the necessary topics is presented at Fig.12 – Fig.13.

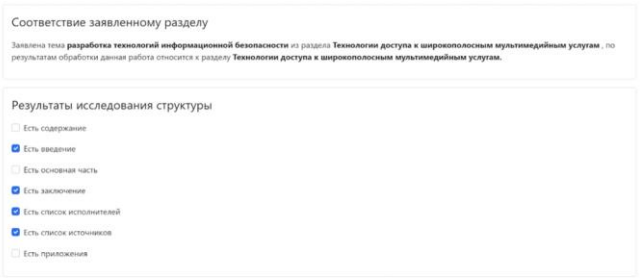

Fig. 12. User interface (screenshot 7)

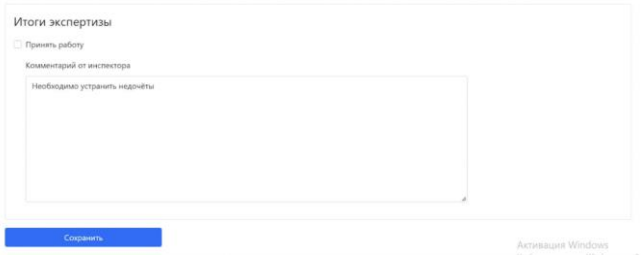

Fig. 13. User interface (screenshot 8)

The results of the report analysis is a graph that shows information about the 30 most frequently used words from the Dictionary of Scientific Terms for R&D reports. The system will also give a result on the presence/absence of the main parts of the report based on this information, the tax officer has to choose in a special window to accept or reject the work submitted for examination.

IV. Conclusion

Due to the growing demand for IT projects and the issuance of exemptions for work done in this area, the tax service in Russia needs to spend a great deal of time manually checking reports and documentation on R&D. That is why there was a need to create a system that automates this process. The paper proposes such a special product for the analysis of textual information intended for the examination of reports received from the taxpayer. have made such a special product for the analysis of text information provided for the examination of reports that come from the taxpayer. The presented functionality allows the tax officer to decide whether or not to accept the work based on the results of the analysis performed by our system, saving time on checking more important parts.

References

[1] Nezhmetdinova Farida Tansykovna, Barabash Natalia Sergeevna Analysis of the management and financing system of interdisciplinary humanitarian research in leading foreign countries (analytical review) // Eurasian Union of Scientists. 2016. No.2-3 (23).

[2] V.F. Sharov Foreign experience in the formation of innovative development mechanisms. – Moscow: Institute of Financial and Economic Research of the Financial University under the Government of the Russian Federation, 2013.

[3] Nalogoviy Kodeks RF - Moscow, Prospect, 2022. – 1168 p.

[4] https://www.consultant.ru/document/cons_doc_LAW_78529/

[5] Anikeychik N.D., Kinzhagulov I.Yu., Fedorov A.V. Planning and management of research and development. Study guide. – St. Petersburg: ITMO University, 2016. – 192 p.

[6] https://www.djangoproject.com/

[7] https://getbootstrap.com/

[8] https://www.postgresql.org/
[9] https://gunicorn.org/
[10] https://www.nginx.com/
[11] https://www.nltk.org/
[12] https://seaborn.pydata.org
[13] https://pandas.pydata.org
[14] https://numpy.org
[15] https://matplotlib.org