



Joint Institute for Nuclear Research

SCIENCE BRINGING NATIONS TOGETHER



MESHCHERYAKOV LABORATORY of INFORMATION TECHNOLOGIES

Analytical platform for intellectual labour market analysis

Belov S. D., Filozova I. A., Gavrilenko Y. E., <u>Ilina A. V.</u>, Javadzade J. N., Kadochnikov I. S., Korenkov V. V., Pelevanyuk I. S., Priakhina D. I., Semenov R. N., Tarabrin V. A., Zrelov P. V.

Introduction

In recent years, the prospects for digital transformation of economic processes have been actively discussed.

This is a rather complex problem that has no solution by traditional methods.

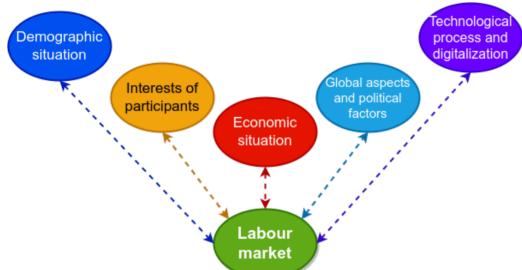
A significant factor for the qualitative development of transformation is the use of Big Data methods and technologies.

One challenge is the use of <u>text mining techniques</u> to assess the needs of regional labor markets in the workforce.

<u>PURPOSE</u>: study of the development of the labor market in the Russian Federation and its relationship with the education system



Relevance



Analysis of the labor market allows you to obtain information about its current condition, make forecasts for a given perspective, assess the level of wages and the degree of compliance with the qualifications and experience of the company's employees, and implement comprehensive measures to motivate employees financially.

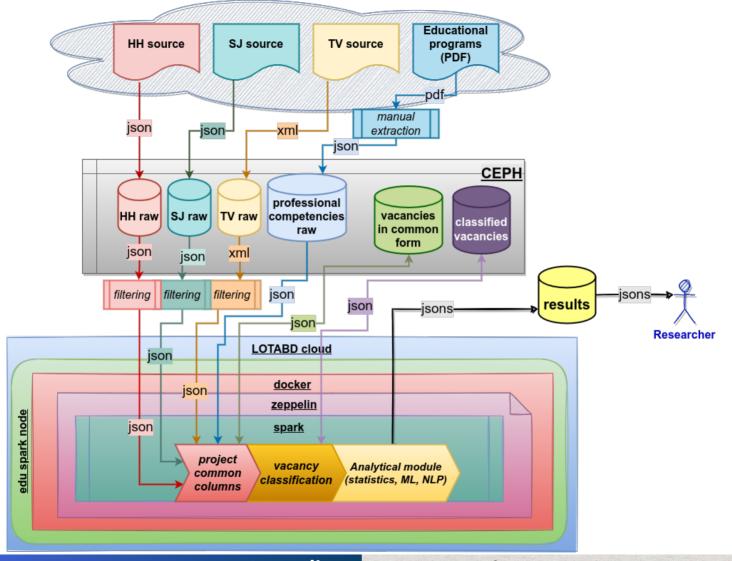
Problems that arise when building an information system for analyzing the labor market:

- Large number of sources of information to be analyzed
- Lack of uniform systematic approaches to the analysis of the labor market in terms of changing the requirements for the qualification of the labor force and reflecting the future needs of the market in the content of educational programs
- Lack of uniformity in the wording of employer vacancies and candidate resumes

Who is the system focused on?

The system is aimed at a wide range of users:

- authorities and the leadership of regions and municipalities;
- management of universities, companies, personnel agencies;
- graduates and prospective students.





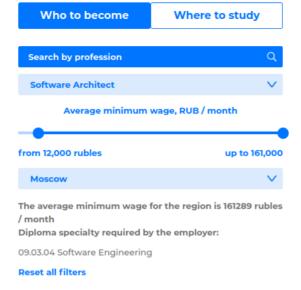
Favourites



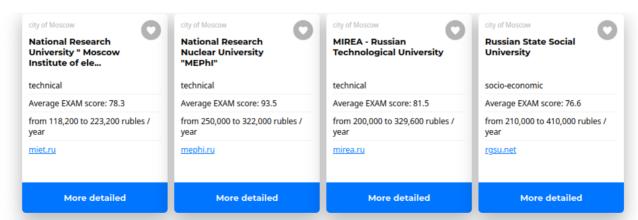




https://na.ria.ru/2022032 3/vuzy-1779514893.html



UNIVERSITIES THAT PROVIDE TRAINING UNDER THESE **PROGRAMS**



Big Data Platform

Trends identification, forecasting, visualization



Modeling



Forecast and prespective analysis



Visualization, services, business analytics, situation centers



Analysis



Semantic models. graph analysis



Machine learning, clustering. classification



Statistical analysis, intellectual analysis, hypothesis testing



Data collection, primary processing and storage



Data collection



Primary processing, filtration



Conversion. compression



Storage

Data sources

External data sources

(Mass media, specialized databases, employment services data, etc.)

Internal materials of organizations

(Educational materials, educational standards, databases, etc.)



SuperJob





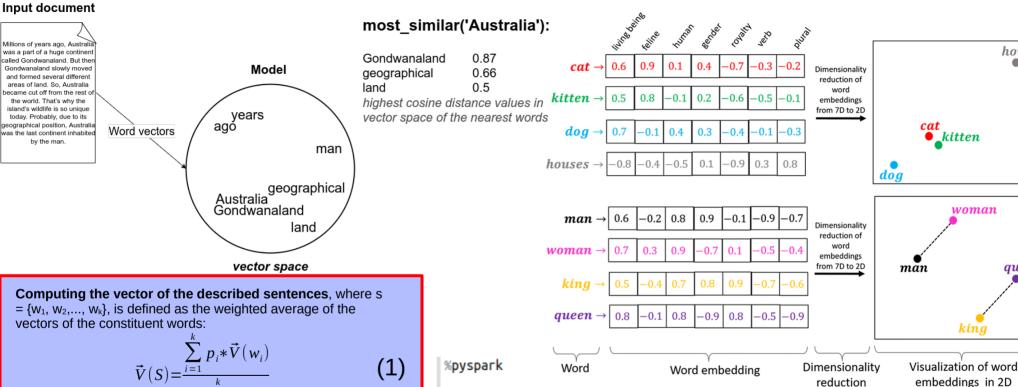






Linking Labor Market requirements to Education supply

(2)



 $V(S) = \frac{\sum_{i=1}^{k} p_i}{\sum_{i=1}^{k} p_i}$ For two words we and we represented as vectors 1) and 2)

For two words w_1 and w_2 represented as vectors 1) and 2), the formula for **calculating semantic proximity** is:

$$\cos(\vec{V}(w_1), \vec{V}(w_2)) = \frac{\vec{V}(w_1) \times \vec{V}(w_2)}{\vec{V}(w_1) \cdot \vec{V}(w_2)}$$

array([0.05447346, 0.1552476 , 0.00851933, 0.06356325, 0.02533658, 0.05318144, -0.02671001, 0.03868017, 0.03270752, -0.08417758, 0.04442836, 0.02490429, -0.11349172, 0.03272674, -0.0954586 , 0.09089028, 0.01185345, -0.0856006 , -0.05252686, -0.03958878, ... (300,)

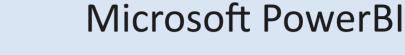
word2vec model['безопасность NOUN']

houses

queen

Two approaches to visualizing results

Technology stack for implementing our own web service

















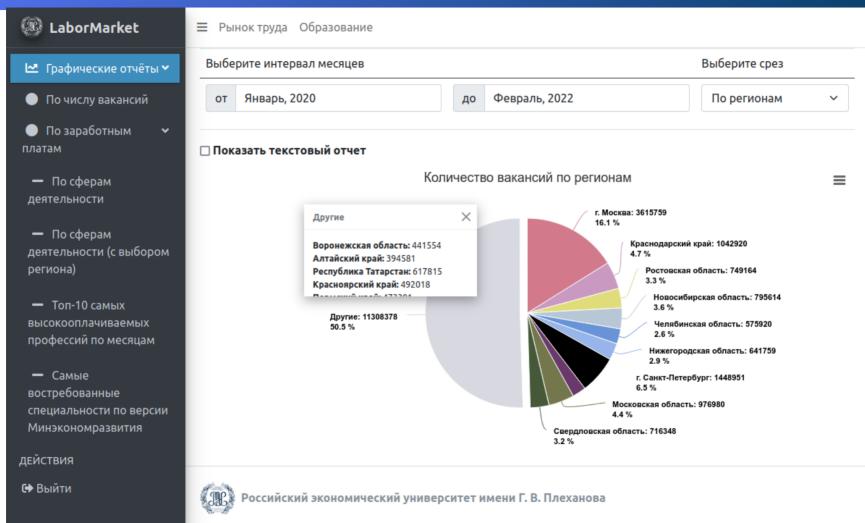




Demonstration of the platform capabilities and some results of the task

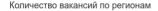
Capabilities of our own web service

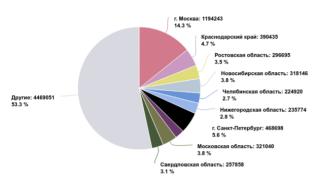
Number of vacancies by region



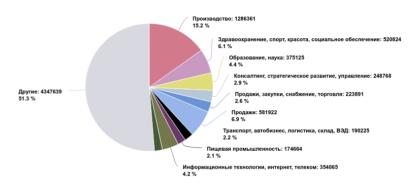
Dynamics of changes in the number of vacancies over two years

2020



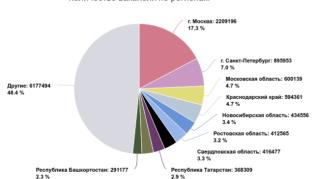


Количество вакансий по проф. областям

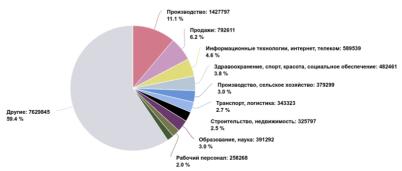


2021

Количество вакансий по регионам

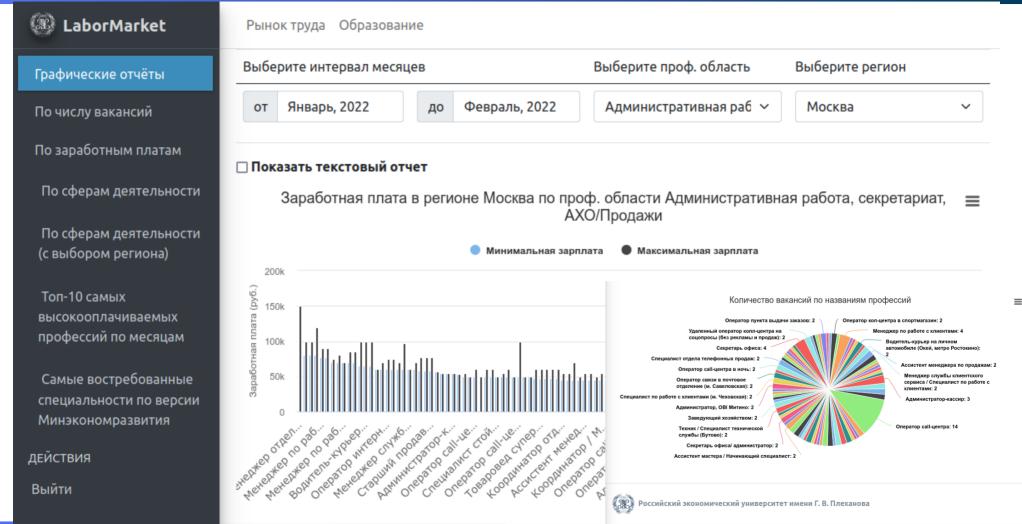


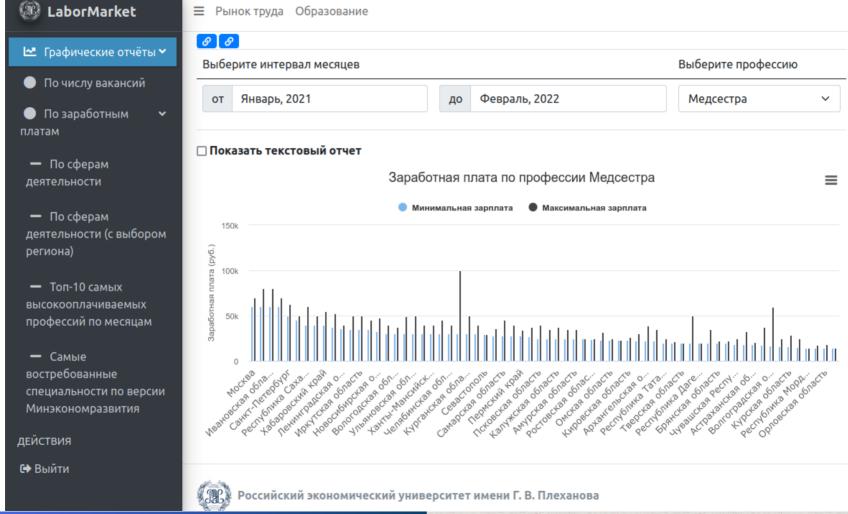
Количество вакансий по проф. областям



Text is not SVG - cannot display

Median salaries by Professional Area and Region (January-February, 2022)

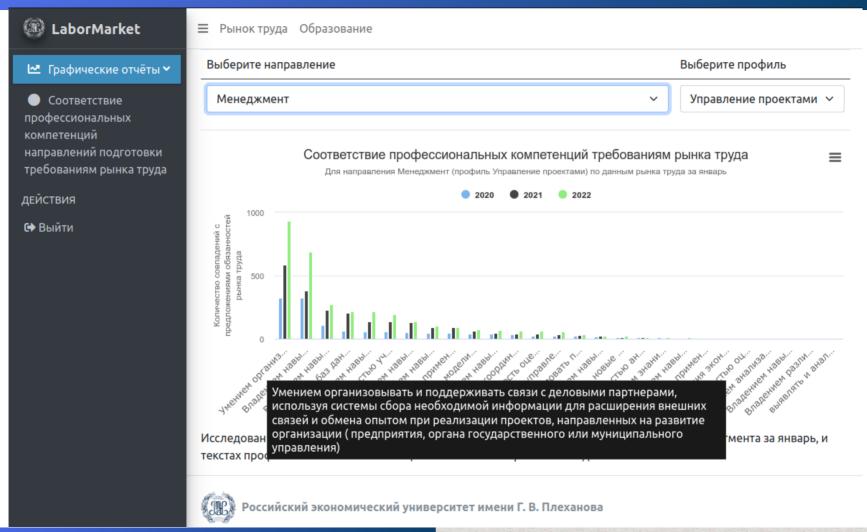




6-8 July, 2022 Anna Ilina MLIT JINR, Dubna

15

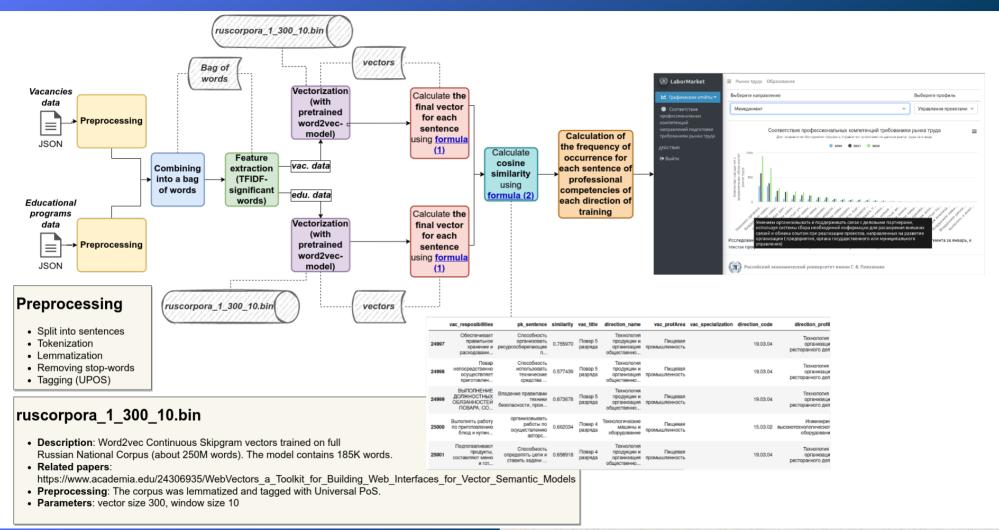
Linking Labor Market requirements to Education supply



6-8 July, 2022 Anna Ilina MLIT JINR, Dubna

16

General flowchart of linking Labor Market requirements and Education System supply



Capabilities of Microsoft PowerBl system

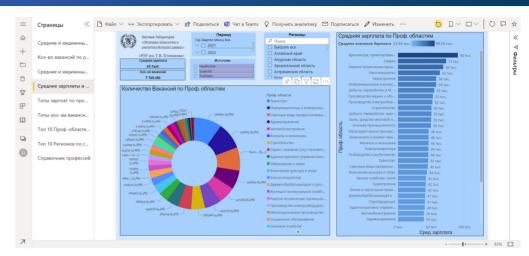
BI — Business Intelligence

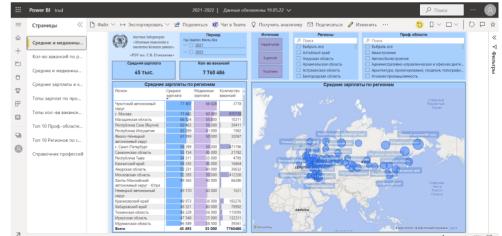
Analyze and visualize data using Business Intelligence

The full set of statistics contains data on the labor market **since 2018**:

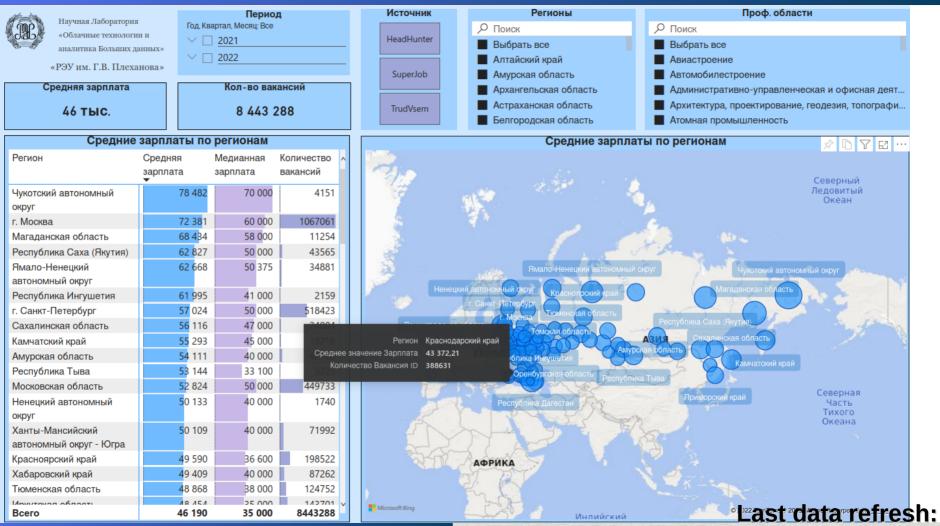
- number of professions (professional standards) studied 1,679,
- the number of vacancies examined is about 20,000,000,
- total raw data more than 10 TB,
- data per month:
- raw <u>hundreds of GB</u>,
- classified less than 10 GB,
- cleaned and prepared for visualization <u>hundreds of MB</u>.







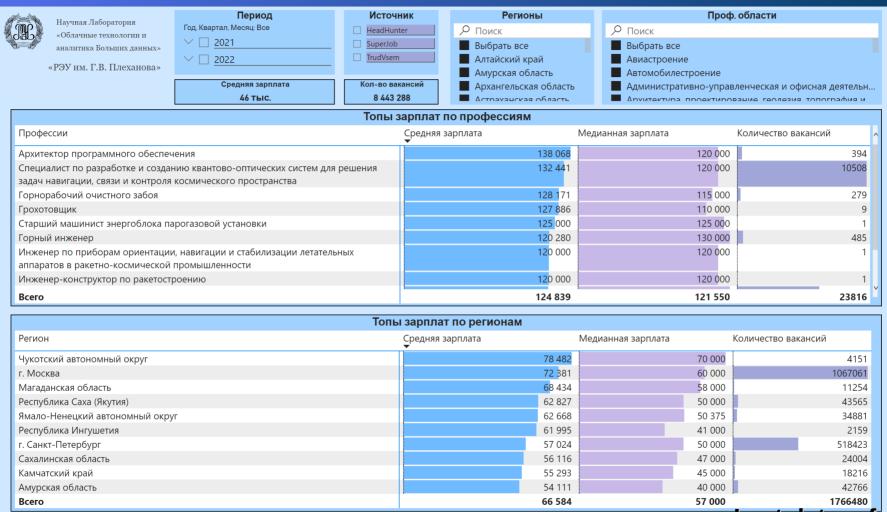
Average and median salaries by region



Anna Ilina

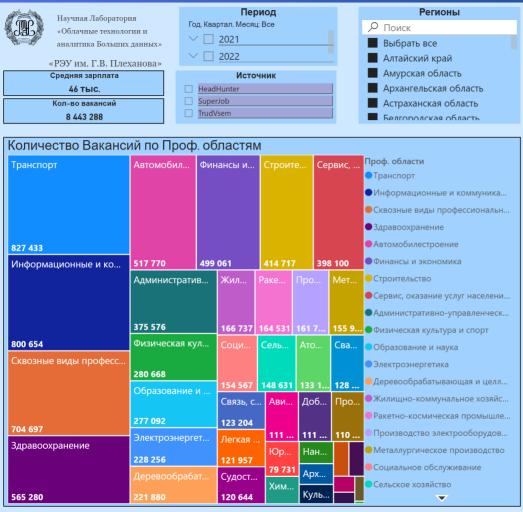
MLIT JINR, Dubna

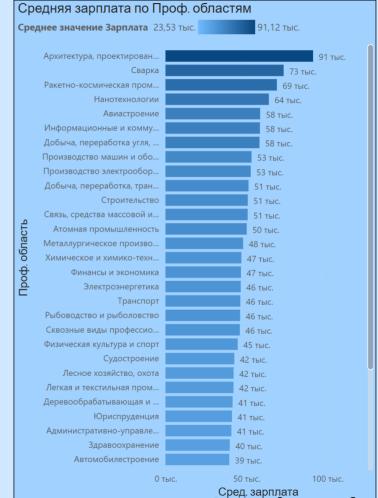
Salary tops by profession and region



Last data refresh:

Average salaries and number of vacancies by professional area





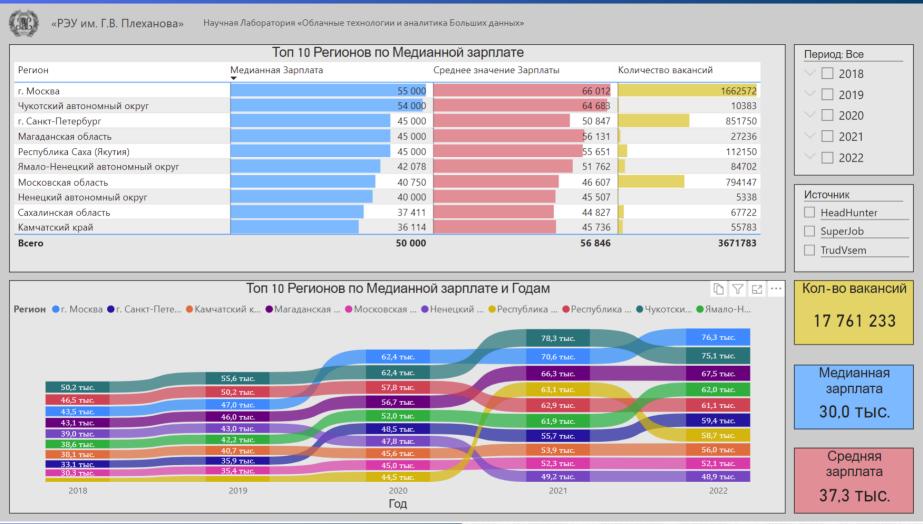
Last data refresh:

Salaries and vacancies in the specialties of the Federal State Educational Standards



Last data refresh:

Top 10 regions by median salary for the year



6-8 July, 2022 Anna Ilina MLIT JINR, Dubna

24

Conclusion

As a result of work, on the basis of the Big Data platform a system was created to study the development of the labor market, including the search for answers to questions related to the response of the education system to this development.

The development of the platform is planned in the following areas:

- Development of the prototype platform to the level of an expert and analytical system. Development of specialized user interfaces (including based on Business Intelligence technologies).
- Expanding the functionality of the platform and the knowledge accumulated in it to take into account additional factors, such as a shortage of personnel, level of training, work experience, etc., as well as identifying hidden patterns.
- Solving the problems of forecasting the situation on the labor market in the constituent entities of the Russian Federation and municipalities and various sectors of the economy.

Thank you for your attention!

26