Equilibrium of Labor Market: New Security Instruments in the Context of Digitalization

Alena Vankevich, *Iryna* Kalinouskaya[*], *Olga* Zaitseva, and *Alena* Korabava

Vitebsk State Technological University, 210038 Vitebsk, Belarus

Abstract. The actual methods of labor market analysis are based on outdated technologies for collecting information and do not consider the competencies available in the CV and in demand by vacancies. In order to obtain reliable current information on the balance of the workforce quality, as the carrier of certain competencies, and market requirements, the method is proposed for determining the degree of their consistency through the ratio of competencies available to applicants and those requested by employers. The proposed methodology, based on big data technologies, uses artificial intelligence as the main toolkit which makes it possible to quickly and efficiently collect, process and visualize the obtained data, which makes it possible to conduct its further qualitative analysis in the context of the proposed / demanded competencies, professions, regions and types of economic activities. As a practical interpretation of the proposed methodology, the paper analyzes the degree of consistency of existing / demanded competencies in the context of the regions of Belarus.

1 Introduction

One of the most important challenges of the modern labor economy is to achieve equilibrium in the labor market. The equilibrium state is achieved when the demand and supply of labor match, it happens when everyone who is looking for a job could find it. The balance between the supply and the demand in the labor market is a set of conformance points, including the conformance between the skills and qualifications required in the labor market and those that job seekers have.

The qualification mismatch issues, as well as their solutions, have become especially relevant in the context of the development of information technologies and universal digitalization, as evidenced by the interest in this issue of scientists, researchers, scientific and international organizations [1, 2, 3, 4]. "Matching skills and jobs are gaining importance with rising unemployment and increasing difficulties for people entering the labour market to find jobs matching their potential. At the same time, structural trends challenge Europe's ability to provide the skills that its economies need to grow and prosper" [5].

---

[*] Corresponding author: i-kalinovskaya@yandex.by

The emergence of "Big Data" has allowed a deeper look at the mismatch issues between the demand for skills and their supply. The emergence and development of artificial intelligence technologies enable the better assess the labor market status [6, 7, 8, 9, 10].

The issues of qualification mismatch are also relevant for the Republic of Belarus. An important current task is to study the supply and the demand in the labor market of the Republic of Belarus using modern methods of processing "Big Data", including artificial intelligence technologies as well as identifying their features for taking further management decisions [11, 12].

The applicable classical methods of labor market analysis do not consider the availability / absence of competencies requested by employers, but consider the demand as the need of the economy for a certain number of workers for a given period of time missing its qualitative aspect [13, 14]. Thus, it became necessary to develop a qualitative approach in analyzing the competence conformance required by employers and applicants indicated in the CV which will allow to consider the qualitative state of equilibrium in the labor market as a balance of the system of available / requested competencies, achieved through training in skills and competencies in demand.

2 Methods for obtaining digital data on labor market using Big Data

The proposed methodology involves the use of "Big Data" technologies for scraping CVs and vacancies from online job search portals in order to group professions and skills according to common European classifiers to identify new skills, the most / least demanded and the most / least represented skills in the CV establishing the degree of conformance (or imbalance) between them [11, 15, 16, 17].

As tools for extracting the required digital information, software products based on artificial intelligence have been developed meeting the speed and volume of processed information and allowing to obtain high-precision results. The architecture of analytical platform for extracting, processing and analyzing data from resumes and vacancies is shown in Figure 1.
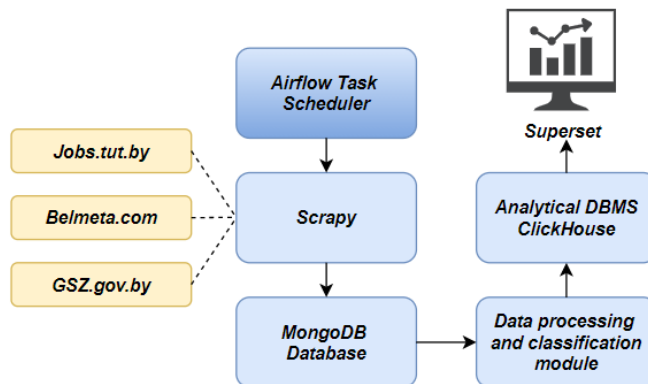
Fig. 1. Architecture of analytical platform for extracting, processing and analyzing data for labor market analysis

The analysis of competencies offered and demanded in the labor market includes the following stages:
1.　Extract of CVs and vacancies from job search websites and other web-sources. In order to identify the optimal data sources containing information about positions

and competencies, an analysis of job search sites [18] was carried out and among them the priority ones were selected – Jobs.tut.by, Belmeta and GSZ.gov.by. To extract CVs and vacancies from selected sources, software modules have been developed based on the Scrapy framework allowing the collection and primary processing of unstructured data in automatic mode.

2. Processing of text data extracted from CVs and vacancies. Since the received data is not structured and contains a large amount of unnecessary information, they are cleaned and brought to a single form. This process included tokenization, stopword and punctuation removal, normalization, stemming, and lemmatization.

3. Deduplication of data in order to remove duplicate resumes and vacancies extracted from different sources using artificial intelligence technologies that allow calculating the similarity of text documents by their vector representation [12, 19].

4. Extraction and classification of positions from CVs and vacancies according to ISCO (International Standard Classification of Occupations). Information from CVs and vacancies was brought to a single form for ISCO job groups. For this purpose, the groups of this classifier were loaded into the system in the form of a vector representation of words and phrases according to which the system searched for the most suitable codes based on their title and description. To form the vector representation, textual models of artificial intelligence (LDA, Word2Vec, ELMo, USE) were considered and evaluated on pre-marked data from CVs and vacancies. As a result, USE (Universal Sentence Encoder) was chosen as the optimal model, which allows you to support various languages and encode whole sentences, which increased the accuracy of processing and comparing text data

5. Extraction and classification of competencies from CVs and vacancies by ESCO using the methods described in the previous clause, while the European classifier was used as a classifier ESCO (European Skills / Competencies, Qualifications and Occupations).

6. Visualization of prepared data from CVs and vacancies on the analytical system platform Superset.

7. Establishment of the degree of consistency between job requirements and applicant CV data.

The proposed methodology allows to quickly and efficiently collect and analyze large data sets of the supply and the demand of the labor market and visually present information of interest in the context of the proposed/demanded competencies, professions, regions and types of economic activity.

3 Analysis of competencies described in CVs and vacancies, their comparison by regions and nature of business

The empirical study base is data from vacancies and resumes, obtained by scraping from the leading online job search portals in Belarus, and prepared according to the proposed method. In the course of study, an analysis of the data array was carried out, including: 227.1 thousand resumes, 16.42 thousand vacancies, 1.57 million CV competencies and 100 thousand vacancy competencies. Information on the distribution of processed data by regions of Belarus is presented in Table 1.

Table 1. Empirical study base.

Indicator	Region						
	Minsk	**Minsk region**	**Mogilev and Mogilev region**	**Brest and Brest region**	**Gomel and Gomel region**	**Grodno and Grodno region**	**Vitebsk and Vitebsk region**
CVs processed	61300	20900	25400	35000	33000	23900	27600
Vacancies processed	11700	908	712	916	777	771	636
CV competencies studied	485000	128000	162000	234000	222000	163000	176300
Vacancy competencies studied	66400	4010	3620	4730	4170	3810	13230

One of the newer and more accurate category of supply and demand analysis in the labor market is to establish consistency between the skills required in vacancies and the skills presented in the resume of applicants. This approach complements the currently used indicator of labor market conditions, which is modified as the ratio of the CV number to the number of vacancies by profession, nature of business and region, or the number of unemployed to the number of vacancies. However, unlike the accepted approaches, it enables analyzing the skill market demand which is more informative since many professions require cross-cutting skills, and emerging professions contain a new combination of skills that were previously singular in more traditional professions.

It is advisable to assess the degree of consistency of skills presented in vacancies and in CVs of applicants using the method of correlation analysis.

Since the compared quantitative data differ in their order, they were ranked before calculating the correlation coefficient. Thus, for the considered ordinal scales, the method of nonparametric analysis was applied and the Spearman's rank correlation coefficient was calculated, which makes it possible to calculate the rank correlation coefficient and does not require the presence of a normal distribution law of correlated series. The results of calculating the correlation coefficient are shown in Table 2 (the significance level of the coefficient is 0.05).

Table 2. Spearman's rank correlation coefficient calculation by ISCO position groups from CVs and vacancies in comparison by Belarus regions.

ISCO group	**Region**						
	Minsk	**Minsk region**	**Mogilev and Mogilev region**	**Brest and Brest region**	**Gomel and Gomel region**	**Grodno and Grodno region**	**Vitebsk and Vitebsk region**
1. Managers	0.956	0.792	0.865	0.855	0.845	0.782	0.729
2. Professi-onals	0.966	0.899	0.925	0.896	0.866	0.885	0.809
3. Techni-cians and associate professionals	0.902	0.754	0.662	0.739	0.664	0.623	0.629
4. Clerical support workers	0.881	0.764	0.764	0.766	0.729	0.708	0.551
5. Service and sales workers	0.943	0.722	0.869	0.844	0.881	0.868	0.644
6. Skilled agricultural, forestry and fishery workers	- *	- *	- *	-0.973	-0.057	-0.168	-0.351

Table 2. Continued

7. Craft and related trades workers	0.909	0.675	0.690	0.731	0.710	0.724	0.619
8. Plant and machine operators and assemblers	0.838	0.657	0.467	0.639	0.551	0.602	0.508
9. Elementary occupations	0.839	0.649	0.576	0.589	0.613	0.627	0.504

[*] - The analysis is impossible due to the absence of described competencies in vacancies and CVs.

The analysis of the obtained correlation coefficients by ISCO position groups has allowed to establish that:

- The <u>high</u> direct relationship (according to the Chaddock scale) in the context of all regions of Belarus has the competence of CVs and vacancies by groups of professions "Managers", "Professionals";
- The <u>noticeable</u> direct relationship has competencies by groups of professions "Technicians and associate professionals", "Clerical support workers", "Service and sales workers", "Craft and related trades workers" and "Elementary occupations";
- The <u>moderate</u> direct relationship has competencies by groups of professions "Plant and machine operators and assemblers";
- The inverse relationship has competencies by groups of professions "Skilled agricultural, forestry and fishery workers" provided that the strength of their relationship differs depending on the region.
- The analysis of the obtained correlation coefficients by regions of Belarus has allowed to establish that:
- Competencies of CVs and vacancies by city of Minsk have the <u>high</u>-strength correlation relationship by all ISCO position groups (without group of professions "Skilled agricultural, forestry and fishery workers");
- Competencies by Minsk, Brest, Gomel, Grodno and Vitebsk regions have <u>noticeable</u> strength of correlation relationship, excluding the group of professions "Skilled agricultural, forestry and fishery workers";
- Competencies by Mogilev region have the <u>moderate</u> strength of correlation relationship.

Provided that the hypothesis of the equality of Spearman's rank correlation coefficient to zero was rejected for all ISCO position groups except for "Skilled agricultural, forestry and fishery workers" in Grodno and Gomel regions. Thus, the rank correlation coefficient of the studied factors is statistically significant, and the rank correlation is significant (with the exception of the indicated regions for the Skilled agricultural, forestry and fishery workers group).

The study conducted allows us to conclude that:

- only in Minsk the structure of skills and the level of qualifications of applicants meets the needs of employers;
- the level of qualification conformance in other regions was distributed as follows (in descending order): Brest and Brest region - Minsk region - Grodno and Grodno region - Gomel and Gomel region - Mogilev and Mogilev region - Vitebsk and Vitebsk region;
- the revealed inverse relationship between the competencies of CVs and vacancies in the Skilled agricultural, forestry and fishery workers group indicates the complete absence of the candidates' knowledge, skills and abilities that are essential

for employers (Brest region), or the partial presence of such knowledge, but those that are optional in the resume of candidates (Gomel, Grodno and Vitebsk regions);

- the lack of the necessary data to analyze the consistency of competencies in the Skilled agricultural, forestry and fishery workers group in the Minsk and Mogilev regions confirms the need to create a unified database of vacancies and resumes describing skills and competencies [11].

4 Conclusion

1. In the context of the economy transformation, it is required to consider the state of equilibrium in the labor market as the balance between the existing/demanded competencies.
2. In order to collect and analyze information about the current state of supply and demand in the labor market, it is advisable to collect vacancies and CVs' competencies from online job search portals using the scraping method. In order to solve this issue, it is legitimate to offer an analytical platform for collecting data for analyzing the labor market.
3. In order to analyze the collected digital information, the article proposes a methodology for collecting the required and available professional and qualification skills, classifying them and establishing the degree of consistency of competencies available to applicants and those demanded by employers, using Big Data and artificial intelligence technologies as tools for data collection and analysis.
4. The systematization of competencies according to the unified European classifiers ISCO and ESCO using artificial intelligence technologies allows to analyze the degree of consistency of job requirements with the data from the CV and to establish missing candidate skills but demanded by the labor market.
5. The comparison of demanded and missing competencies in the context of regions and nature of business allows to develop measures to improve the efficiency of educational processes and state policy in the field of employment with reference to the specific region.

Acknowledgements

References

1. A. Sakamoto, J. Sung, Skills and the Future of Work, ILO (2018)
2. R. Almeida, J. Behrman, D. Robalino, *Rethinking Training Policies for Workers* (2012)
3. *The future of work and skills in ETF partner countries*, ETF Issues Paper (2019)
4. Cedefop, Developing and matching skills in the online platform economy: findings on new forms of digital work and learning from Cedefop's CrowdLearn study, Cedefop reference series, 116 (2020)
5. Cedefop, The skill matching challenge (2010)

6. E. Colombo, F. Mercorio, M. Mezzanzanica, Information Economics and Policy, **47**, 27 (2019)

7. S. D. Belov, J. N. Javadzade, I. S. Kadochnikov, V. V. Korenkov, P. V. Zrelov NEC'2019, 469 (2019)

8. Cedefop, The online job vacancy market in the EU: driving forces and emerging trends, Cedefop research paper, **72**, 96 (2019)

9. M. Feldman, F. Juldaschewa, A. Bernstein, *Data Analytics on Labor Markets* (2017) https://arxiv.org/

10. J. J. Horton, P. Tambe, Big Data, **3**, 130 (2015)

11. A. Vankevich, I. Kalinouskaya, E3S Web of Conferences, **208**, 03060 (2020)

12. A. Vankevich, I. Kalinouskaya , Belarusian Economic Journal, **2(91),** 38 (2020)

13. G. J. Borjas, Labor Economics, **7**, 8 (2016)

14. P. Ehrenberg, P. C. Smith, *Modern economy of labour: The theory and state policy*, 800 (1996)

15. L. M. Kureková, M. Beblavý, A.Thum-Thysen, Labor Econ **4,** 18 (2015)

16. R. Strack, E. Kaufman, A. Kotsis, M. Sigelman, D. Restuccia, B. Taska, BCG Burning glass (2019)

17. The feasibility of using big data in anticipating and matching skills needs – International Labour Office, Geneva: ILO (2020)

18. E.V. Vankevich, O.V. Zaitseva, Digital transformation, **2(11)**, 33 (2020)

19. M. Mezzanzanica, F. Mercorio, *Big Data for Labour Market Information*, A Brief Introductory Guide (2019)