

METHODOLOGY FOR THE MINING OF HUMAN RESOURCES IN THE DIGITAL ECONOMY**Ivanov S., V. Maltuz V., Kaliuzhna Iu., Terentieva N.***Zaporizhzhia National University
Ukraine, 69600, Zaporizhzhia, Zhukovsky str., 66*flydaiver@gmail.com, ma.viktorija@ukr.net,
kalyuzhnaya.ju@gmail.com, terenteva_nataliya@ukr.netORCID: 0000-0003-1086-0701, ORCID: 0000-0002-3863-6456,
ORCID: 0000-0002-3335-6551, ORCID: 0000-0001-6930-879X**Key words:**

human resources, Big Data, Industry 4.0, data-mining model, fuzzy modelling, digital marketing, digital body, forecasting, human capitals.

The In this article, a systematic methodology for analyzing and assessing the effectiveness of human resources based on fuzzy sets using big data technologies is used. Based on our research, we analyzed the big data construction method for our chosen approach using Industry 4.0.

For the selected fuzzy sets, a set of sequence of procedures in the sequence of the method for assessing the effectiveness of human resources have been identified. Input and output membership functions for data mining have been developed.

This article discusses process of building rules of fuzzy logic that allowed us to determine the degree of truth for each condition.

The relevance achieved through the development of a methodology that includes eight procedures required for a comprehensive assessment of the economic efficiency of human resources. In this article, an approach to assessing the normative or average values of the performance of official duties by employees of an enterprise in many specialties, educational levels, levels of management, as well as taking into account the description of many positions, descriptions of compliance and interchangeability of positions, assessment of additional characteristics of employees and a description of many additional tasks and their characteristics is presented. The article presents a structural data-mining model for personnel assessment. The results of modeling the assessment of human resources is presented.

МЕТОДОЛОГІЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЛЮДСЬКИХ РЕСУРСІВ В ЦИФРОВІЙ ЕКОНОМІЦІ**Іванов С., Малтюз В., Калюжна Ю., Терент'єва Н.***Запорізький національний університет
Україна, 69600, м. Запоріжжя, вул. Жуковського, 66***Ключові слова:**

людські ресурси, Великі Дані, Індустрія 4.0, модель інтелектуального аналізу даних, нечітке моделювання, цифровий маркетинг, цифрове тіло, прогнозування, людський капітал.

У статті використовується систематизована методологія аналізу та оцінки ефективності людських ресурсів на основі нечітких множин із використанням технологій великих даних. На основі нашого дослідження ми проаналізували метод побудови великих даних для обраного нами підходу, використовуючи Індустрію 4.0.

Для відібраних нечітких множин, набір послідовності процедур в послідовності методу для оцінки ефективності людських ресурсів були ідентифіковані. Розроблено функції вводу та виведення для приєднання даних. Ця стаття досліджує процес правил нечіткої логіки, що дозволяє нам визначити ступінь правдивості для кожної умови.

Актуальність дослідження полягає у розробці методології, яка включає вісім процедур, необхідних для всебічної оцінки економічної ефективності людських ресурсів. У цій статті запропоновано підхід до оцінки нормативних або середніх значень виконання службових обов'язків працівниками підприємства за багатьма спеціальностями, освітніми рівнями, рівнями управління, а також з урахуванням опису багатьох посад, описів відповідності представлена взаємозамінність посад, оцінка додаткових характеристик працівників та опис багатьох додаткових завдань та їх характеристик.

У статті представлена структурна модель аналізу даних для оцінки персоналу. Представлено результати моделювання оцінки людських ресурсів.

General problem resolution

Today business is forced to solve a whole range of complex and unique tasks. To solve such problems as the tasks of increasing and stabilizing the development of

economic objects in modern conditions of economic activity, new approaches are required, which determined the emergence of a new concept by the German economist Klaus Schwab, President of the World Economic Forum. Forum. Forum in Davos [1]. According to this concept, it

is argued that we live in the era of the fourth industrial revolution, when the virtual world is combined with the physical world using information technology. The fourth industrial revolution is characterized by a change in economic relations and the widespread use of intelligent technologies (cloud technologies, big data, artificial neural networks and fuzzy sets, data mining, and others). These technologies form the backbone of the digital economy.

Therefore, for the successful development of states, regions, enterprises in the era of the fourth industrial revolution (Industry 4.0), an assessment of the possible potential efficiency of human resources (HR) plays an important role, since HR can radically change and increase the development and competitiveness of business entities in modern conditions. In addition, the global COVID-19 pandemic has forced countries to rethink their national concepts of economic development.

Beyond these areas, it should be noted that for a long time HR management has been focused on standardization and versatility. However, today this approach is gradually becoming obsolete. This approach is being replaced by methods of personnel management focused on the maximum use of the intellectual capital of employees.

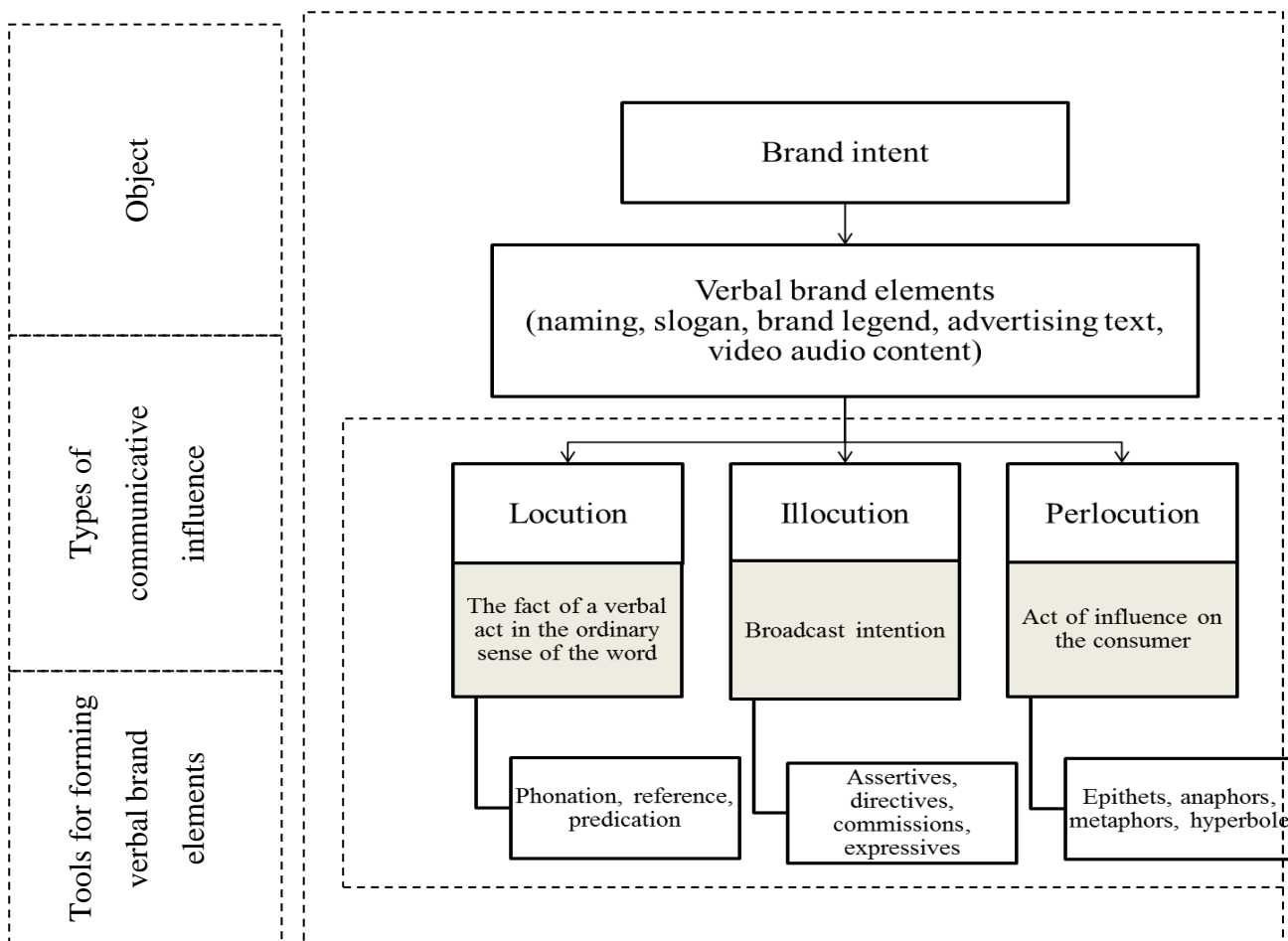
Analysis of last researches and publications

At the present stage, the financial crisis is one of the Ensuring compliance with the modern requirements of Industry 4.0 and available technologies for using various technologies for processing Big Data is based on the use of previously developed directions.

In the course of the recent work of M. Ivanov and S. Ivanov [6], we realized that there is a great need to develop a method for representing big data HR. It is also necessary to develop a methodology for analyzing and assessing the economic efficiency of personnel based on big data for the development of a region or enterprise.

The problems of rating management have not yet been resolved. The solution to the problem of rating management was the work of Yu. Lysenko, V. Petrenko, O. Bogatov [7]. However, the level of personnel development and their assessment were not taken into account. V. Helman [8] considers the development of enterprise personnel as a change in its qualitative characteristics, in which indicators are offered in the form of a degree of activity.

In the work of Brian E. Becker [9] approaches are considered in which all managers should be involved in the management process, where the role of employees is important for the competitive advantage of the enterprise. The results of the work of scientists Mark A. Huzelid [10], Jeffrey Pfeffer and John F. Weighi [11] show that successful enterprises have several characteristics in common: stable job security, high levels of self-government and excellent wages. Dennis R. Briscoe [12] in his work writes about a situation when each employee of an enterprise effectively performs his duties and builds a highly efficient work system in which the employee bears maximum involvement and responsibility.



In modern enterprises, balancing the need to coordinate and synchronize HR across cities and around the world is an important challenge, as discussed by Randall S. Schuler, Pavan S. Budhwar and Gary W. Florkowski [13]. Achieving this balance is becoming increasingly difficult due to the level of functional diversity that states, regions and enterprises strive for in the era of Industry 4.0. Approaches to assessing the performance indicators of personnel in the context of the development of the digital economy are considered in the work of S. Ivanov [14].

The use of big data in human resource management has been reviewed by P. O'Donovan [15] and A. Caggiano [16].

Along with these studies, not enough attention is paid to the problems of fuzzy modeling of big data of human resources in the conditions of Industry 4.0, which makes this problem very relevant today.

The purpose of the article

The article is devoted to intellectual analysis and assessment of the effectiveness of human resources in the context of Industry 4.0. In addition, our goal was to reveal new knowledge based on the application of fuzzy set

theory with the possibility of using Big Data.

Study results

Modern construction and architecture of Big Data allows not only storing, but also processing and analyzing data that is too large or too complex for traditional database management systems.

In Industry 4.0, according to research [3], they are required by using a strategic change management approach that gives them a broader approach to leverage the benefits of analytic processing.

Big Data offers tremendous opportunities to revolutionize human resource management.

Managing human resources with cloud-based solutions opens up new opportunities and solutions.

These solutions include a new level of accessibility that facilitates greater employee mobility. The ability to effectively apply data mining tools and decision-making systems. Great opportunities, flexibility and constant updating of Big Data contribute to the development of theoretical and practical developments. The Big Data architecture can be represented as follows (Fig. 1).

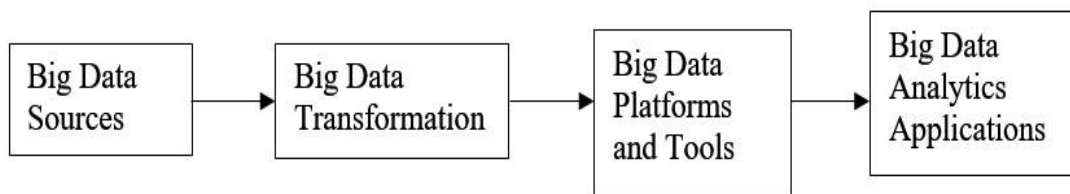


Figure 1 - Architecture Big Data analytics.

Today, the sources of Big Data for us are the Internet (social networks, websites and other applications). Big Data transformation is based on Data Warehouse principles.

The standard description of a multidimensional data model is based on the following concepts: Data Hypercube, Dimension, Memders, Cells, and Measure.

A data hypercube contains one or more dimensions and is an ordered collection of components. Each component is defined by one and only one set of measurement values - labels. The component can contain a measure or be empty.

A dimension is understood as a set of marks that form one of the faces of a hypercube. An example of a time dimension is a time period: day, month, quarter, year. An example of an economic dimension is a list of profitability indicators: working capital, cost, etc.

To gain access to the data, we need to specify one or more directions for choosing the measurement values that correspond to the necessary components. The mechanism for selecting measurement values is the fixation of marks, and the set of selected measurement values is a set of fixed marks.

The set of dimensions of a hypercube can be written in the following form

$$P = \{p_i = (op1, op2, ep1, ep2, ep3)\}, i = (1, N)^- \quad (1)$$

where op1 – is a generalized indicator of job compliance, characterizing the degree of conformity of qualifications

and work experience of the post, level of responsibility, as well as the quality of the performance of current work and duties,

op2 – is a generalized indicator of diligence, characterizing the effectiveness of the tasks (complexity, quality, timeliness),

ep1 – ambitiousness, a single indicator of personality characteristics,

ep1 – the quality of a leader, an indicator of personality characteristics,

ep3 – the level of attitude in the team, a single indicator of personality characteristics.

Then the set of measurement labels p_i is written

$$M(p_i) = \{m_1, m_2, \dots, m_n\}, i = (1, n)^- \quad (2)$$

And the set of fixed dimensions $D^A = \{p_1^A, p_2^A, \dots, p_n^A\}$ and fixed labels of the fixed dimension of the hypercube

$$M(p_i^A) = \{m_1^A, m_2^A, \dots, m_n^A\}, i = (1, n)^- \quad (3)$$

The data hypercube will be denoted as the set of its cells $H(D, M)$, which corresponds to the sets D, M . Then the subset of the data hypercube HR , will correspond to the set of fixed values D^A, M^A and we will denote it as $H^A(D^A, M^A)$.

Thus, we have built a method for creating a visual representation of a multidimensional database. This allowed us to assess the effectiveness of personnel taking into account the filling of data hypercubes and to carry out a visual search for information in the database.

The task of fuzzy modeling and data mining when managing HR processes is to efficiently extract and analyze the existing data array of employees with subsequent management of personnel using cloud solutions. This will allow the rapid

implementation of a new personnel management system, obtaining a new level of accessibility and increasing its mobility.

The resulting performance indicators of personnel at the enterprise can be represented in the form of multidimensional structures, where the corresponding indicators of the enterprise management system represent each measurement. The following method of modeling human resources is propose, which is presented in Fig. 3.

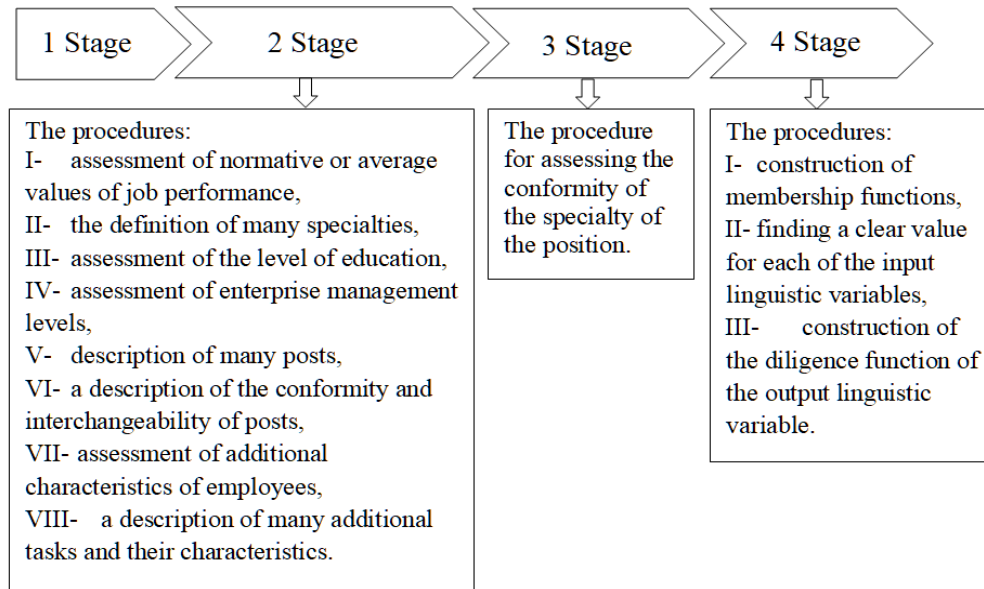


Figure 3 -The method of modelling human resources in fuzzy management.

The proposed method includes four stages.

Stage 1 Procedures. The first stage solves the problem of choosing the analyzed indicators. For this, a lot of ratings are determined (1).

Stage 2 Procedures. At the second stage, the initial information is determined, which is necessary for calculating indicators based on expert assessments, analytical indicators (for example, work experience, quality of work performed, and others).

To describe the formalized set of sets of source information, we introduce the rules, namely, if the set $P = \{p_i = (op1, op2, op3)\}$, is defined, then to use the value of the component $op2$ of unit level 0, we will use the notation $op2$ (join operator).

At the second stage, procedures are applied that allow: The first (I) procedure allows you to evaluate the regulatory or average value of the performance of official duties by employees – $P0$:

$$P0 = \{p0_i = (op01, op02)\}, i = \overline{1, N}, \quad (1)$$

where $op01$ – normative or average value of the job performance of the i -th employee,

$op02$ – normative or average value of the level of assessment of the performance of tasks of the i -th employee.

The second (II) procedure is aimed at identifying many specialties (economist, programmer, builder and others) – SP :

$$SP = \{sp_r\}, r = \overline{1, L^{sp}}, \quad (2)$$

where sp_r – is the r -th specialty,

L^{sp} – is the number of specialties.

The third procedure allows you to assess the level of education (secondary, bachelor, master and others) – UO : $UO = \{uo_v = (name, \mu)\}, v = \overline{1, L^{uo}}, \quad (3)$

where uo_v – is the vector of characteristics of the v -th category,

$name$ – category name,

μ – assessment of the level of education for the category in points,

L^{uo} – is the number of categories.

The fourth (IV) procedure is aimed at assessing the level of enterprise management (higher, middle and lower level) – UD :

$$UD = \{ud_w = (name, \gamma)\}, w = \overline{1, L^{ud}}, \quad (4)$$

where ud_w – is the vector of characteristics of the w -th level,

$name$ – level name,

γ – is an estimate of the level in points,

L^{ud} – the number of levels, which is determined by the scale of the enterprise.

The fifth (V) procedure solves the problem of describing many positions in the enterprise – D :

$$D = \{d_j = (name, ud, uo, kl, SPD): ud \in UD, uo \in UO\}, j = \overline{1, L^d}, \quad (5)$$

$$SPD = \{spd_r = (sp, \beta): sp \in SP, 0 \leq \beta \leq 1\}, r = \overline{1, L_j^{sp^d}}, \quad (6)$$

where d_j – is the j -th position,

name – job title,

ud – position level in the organizational and staff structure of the enterprise,

uo – the level of education required for the j -th position,
kl – required work experience (minimum number of years) in a given position for an optimal qualification level,

SPD – many specialties related to this position,

spd_r – is the vector of the correspondence characteristics of the r -th specialty of the j -th position β ,

β – is the correspondence coefficient of the specialty *SP* of the j -th position,

L^d – is the number of posts,

$L_j^{sp^d}$ – number of specialties in the j -th position.

The sixth (VI) procedure solves the tasks of describing correspondence and job interchangeability – *SD*:

$$SD = \{sd_f = (d1, d2, \alpha): d1 \in D, d2 \in D, 0 \leq \alpha \leq 1, (d1 = d2) \Rightarrow \alpha = 1\}, f = \overline{1, L^{sd}}, \quad (7)$$

$$L^{sd} = (L^d)^2, \quad (8)$$

where *sd_f* – is the f -th vector of job matching characteristics *d1* and *d2*,

α – is the compliance coefficient.

The seventh (VII) procedure is aimed at assessing additional characteristics of employees – *A*:

$$A = \{a_i = (ds, ST, OB): ds \in D, uo \in UO\}, i = \overline{1, N}, \quad (9)$$

$$ST = \{st_j = (d, kl): d \in D, d \in UO\}, j = \overline{1, L_1^{st}}, \quad (10)$$

$$OB = \{ob_w = (sp, uo, god): sp \in SP, uo \in UO\}, w = \overline{1, L_1^{ob}}, \quad (11)$$

where a_i – is the vector of characteristics of the i -th employee,

ds – the position held by the employee,

ST – many posts in which the employee previously worked and experience in them,

st_j – vector of characteristics of work experience in previous positions,

kl – length of service (number of years) in the position d ,

OB – value, reflects the education received by the i -th employee,

sp – specialty,

uo – level of education,

god – year of receipt of the qualification document (certificate, certificate, diploma and others),

L_i^{st} – the number of posts previously held by the i -th employee,

L_i^{ob} – the number of specialties in which the employee was educated by the i -th employee.

The eight (VIII) procedure allows you to describe many additional tasks (determined by orders) and their characteristics in the enterprise:

$$Z = \{z_k = (t0, tk, tk', usz), 0 \leq \beta \leq 1\}, k = \overline{1, M}, \quad (12)$$

where Z_k – is the vector of characteristics of the k -th task,
t0 & *tk* – the value of the beginning and end of tasks, determines the term for completing the task in units of measurement of working time (for example, working days, hours and others),

tk' – the value of time, determines the critical deadline for completing the task, after which the task is either canceled or transferred to another performer,

usz – task difficulty level,

M – the number of tasks.

The set of completing additional *IZ* tasks by employees can be written as follows:

$$IZ = \{iz_k = (a, z, uv^p, uv): a \in A, z \in Z, 0 \leq uv^p \leq 200\}, k = \overline{1, M}, \quad (13)$$

where iz_k – is the characteristic of the k -th job,

a – an employee who performs additional tasks,

Z – the task

uv^p – is the percentage of the task according to the plan at the current time t ($uv^p = 0$ at time $t0$, $uv^p = 200$ at time k),

uv – is the percentage of the task at the current time t .

In case of failure to perform additional tasks, the value of $IZ = 0$.

Stage 3 Procedures. At the third stage, the procedure for assessing the conformity of the specialty of the position is performed. The function $f\beta$ returns the value of the correspondence of the specialty *xsp* to the position *xd*:

$$f\beta(xsp, xd) = \begin{cases} d_{j_0} \cdot spd_{i_0} \cdot \beta, \exists j_0, r_0: (d_{j_0} = xd) \wedge (d_{j_0} \cdot spd_{r_0} \cdot sp = xsp) \\ 0, \neg \exists j_0, r_0: (d_{j_0} = xd) \wedge (d_{j_0} \cdot spd_{r_0} \cdot sp = xsp) \end{cases} \quad (14)$$

The function $f\alpha$ returns the value of the coefficient of correspondence and interchangeability of the post *xsp* post *xd*:

$$f\alpha(xds, xd) = \begin{cases} sd_g \alpha, \exists g_0: (sd_{g_0} \cdot d1) \wedge (sd_{g_0} \cdot d2 = xd) \\ 0, \neg \exists g_0: (sd_{g_0} \cdot d1) \wedge (sd_{g_0} \cdot d2 = xd) \end{cases} \quad (15)$$

To determine job conformity is the level of education of the position held in conjunction with work experience in similar or related positions:

$$p_i.op1 = \delta(op11 \cdot op12) \cdot op13, \quad (16)$$

$$op11 = \sum_{w=1}^{L_i^{ob}} f\beta(a_i.ob_w.sp, a_i.ds) \cdot \frac{a_i.ob_w.god}{godT}, \quad (17)$$

$$op12 = \sum_{j=1}^{L_i^{st}} f\delta(a_i.ds.f, a_i.st.d) \cdot \frac{a_i.st.j.kL}{a_i.ds.kL}, \quad (18)$$

where $godT$ – is the value of the current year,

$op11$ – qualification level of education received,

$op12$ – qualification level, which is determined by work experience,

$op13$ – quality of job performance, determined by an expert.

When solving the problem of data mining in the management of HR processes, fuzzy logic methods are used to display the result on the interval [0; 1].

Stage 4 Procedure. Therefore, at the fourth stage, the procedure for constructing membership functions based on the theory of fuzzy sets is performed.

The authors of the article propose a solution to the problem of constructing a data mining methodology in human resource management based on Big Data.

Conclusions

This article proposes a mining methodology for assessing the effectiveness of human resources. A four-stage method for analyzing and evaluating HR effectiveness has been developed. At the first stage, the problem of choosing the analyzed indicators is solved. At the second stage, eight procedures are performed to solve the following tasks. This procedure allows you to obtain an assessment of the standard or average values of labor productivity, to determine a number of specialties, to assess the level of education, to assess the levels of enterprise management, to describe many job responsibilities, to determine the conformity and interchangeability of work, to assess the additional characteristics of employees. At the third stage, the procedure for assessing the conformity of the specialty to the position held is carried out. At the fourth stage, the procedure for constructing membership functions based on the theory of fuzzy sets is performed. The proposed methodology can be used to assess the human resources of cities, regions and the country as a whole.

References

- Schwab, Kl.: The Fourth Industrial Revolution. Currency, Geneva (2016)
- Thoben, K., Wiesner, S., Wuest, T.: "Industrie 4.0" and Smart Manufacturing – A Review of Research Issues and Application Examples. International Journal of Automation Technology. 1, 4-19 (2017)
- Wichmann, R., Eisenbart, B., Kilian Gericke K.: The Direction of Industry: A Literature Review on Industry 4.0. Proceedings of the Design Society. International Conference on Engineering Design ICED19. 1, 2129-2138 (2019). doi:10.1017/dsi.2019.219
- Kuyoro, S., Ibikunle, F., Oludele, A.: Cloud Computing Security Issues and Challenges. International Journal of Computer Networks (IJCN). 3, 247-255 (2011) . [in Ukrainian].
- Witten, I., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Elsevier, San Francisco (2005) . [in Ukrainian].
- Ivanov, M., Maksyshko, N., Ivanov, S., Terentieva, N.: Intelligent Data Analysis in HR Process Management. CEUR Workshop Proceedings. 2608, 754-768 (2020)
- Lysenko, Yu., Petrenko, V., Bogatov, O., Skobelev, V.: Reytingovoye upravleniye ekonomicheskimi sistemami. Yugo-Vostok, Donetsk (1999)
- Helman, V.M., Makazan, Ye.V., Buriak, A.M.: Personnel development as pledge of the success of enterprise. Bulletin Zaporizhzhia national university Economic sciences. 43(3), 112-115 (2019). doi:10.26661/2414-0287-2019-3-43-18
- Becker, Br., Huselid, M.: Strategic Human Resources Management: Where Do We Go from Here?. Journal of Management. 6(32), 898-925 (2006). doi: 10.1177/0149206306293668. . [in Ukrainian].
- Huselid, M.: (1995). The Impact of Human Resource Management Practices on Turnover and Productivity. The Academy of Management Journal. 3(38), 635-672 (1995) . [in Ukrainian].
- Pfeffer, J., Veiga, J.: Putting People First for Organizational Success. Academy of Management Perspectives. 13(2), 37-48 (1999) . [in Ukrainian].
- Briscoe, D. R., Schuler, R. S., Claus, L.: International Human Resource Management: Policies and Practices for Multinational Enterprises. Routledge, New York (2009)
- Schuler, R., Budhwar, P., Florkowski, G.: International human resources management. Handbook of international management research. 2 ed. Michigan University Press, Michigan (2003) . [in Ukrainian].
- Ivanov, M., Ivanov, S., Terentieva, N., Maltiz, V., Kalyuzhnaya, J.: Fuzzy modeling in human resource management. E3S Web of Conferences. 166, 1-8 (2020). doi:10.1051/e3sconf/202016613010. [in Ukrainian].
- O'Donovan, P., Gallagher, C., Leahy, K., O'Sullivan D.: A comparison of fog and cloud computing cyber-physical interfaces for Industry 4.0 real-time embedded machine learning engineering applications. Computers in Industry. 110, 12-35 (2019) . [in Ukrainian].
- Caggiano, A.: Cloud-based manufacturing process monitoring for smart diagnosis services. International Journal of Computer Integrated Manufacturing. 31, 612-623 (2018) . [in Ukrainian].