Fuzzy modelling of Big Data of HR in the conditions of Industry 4.0

Mykola Ivanov[1[0000-0002-1908-0763]], Sergey Ivanov[1[0000-0003-1086-0701]], Olexander Cherep[1[0000-0002-3098-0105]], Nataliia Terentieva[1[0000-0001-6930-879X]], Victoria Maltiz[1[0000-0002-3863-6456]], Iuliia Kaliuzhna[1[0000-0002-3335-6551]] and Vitaliy Lyalyuk[2[0000-0001-7258-2079]]

[1] Zaporizhzhia National University, 66 Zhukovsky Str., Zaporizhzhia, 69063, Ukraine
[2] State University of Economics and Technology,
5 Stepana Tilhy Str., Kryvyi Rih, 50006, Ukraine
nn_iva@ukr.net, flydaiver@gmail.com, cherep2508@gmail.com, terenteva_nataliya@ukr.net, ma.viktorija@ukr.net, kalyuzhnaya.ju@gmail.com, vitalij.lyalyuk@gmail.com

Abstract. 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.

Keywords: human resources, Big Data, Industry 4.0, data-mining model, fuzzy modelling.

1 Introduction

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 in Davos [21]. 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 [13], big data [12], artificial neural networks and fuzzy sets [23], data mining [7], 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 [22].

It should be said that the term Industry 4.0 is relatively new, which is applied to the concept of the fourth industrial revolution and covers a wide range of modern technologies and approaches, mainly related to the digital economy. In terms of modern technologies, Industry 4.0 is associated with the application and intelligent processing of data in the following areas: artificial neural systems, industrial Internet, cloud solutions and decentralized services, as well as processing and mining of large amounts of data. The works of Klaus-Dieter Thoben, Stefan Wiesner and Thorsten Wuest [27], Robert Lawrence Wichmann, Boris Eisenbart and Kilian Gericke [28] are devoted to these studies on the technology of processing big data in real time. According to the application of the approaches and technologies of using Industry 4.0, they can be divided into the following areas: Big Data storage, data mining and the use of artificial neural systems and fuzzy sets. As for the use of big data, first of all, one can highlight the cloud platforms Amazon Web Services (AWS), Google Cloud Platform, Microsoft's Azure and others. These decisions influenced the formation of cloud computing concepts discussed in the work of S. O. Kuyoro, F. Ibikunle and A. Oludele [14], as well as in the work on data mining and machine learning by Ian H. Witten and Elbe Frank [29] and others [11; 18].

The first direction is determined by the level of application of cloud computing. High information technologies are sweeping the world and are replacing classical methods of analysis with HR processes. Nowadays, robotic programs are being used that offer employees of the enterprise to pass an express interview or interview using expert systems, which are still in use.

The next direction is based on the application of data mining in relation to their historical layers.

This analysis and assessment of HR allows you to select employees who are capable of solving modern problems. Now new specialists under 25 are entering the labour market, who have completely different knowledge, interests and fundamentally different ideas about modern work. Young specialists are able to quickly make management decisions and promote new projects, thereby increasing the level of work at the enterprise.

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. This is stimulated by the simultaneous satisfaction of individual needs, desires and capabilities of employees and their synchronization with the tasks of the enterprise. Modern HR specialists are beginning to more closely monitor the development of employees in the region and within the enterprise, which allows flexible management of career growth, which can be adjusted taking into account the proposals of the employees themselves.

2 Research motivation and formal problem statement

In this section, we provide our motivation for this work and provide an analytical overview of the research questions used to develop a method for analytically assessing human resources.

2.1 Research motivation

Our main motivation for this research work is to bridge the gap between regional, production development scenarios, taking into account the assessment of the economic efficiency of human resources using big data. We are confident that this work will serve as a basis for developers of use cases for Industry 4.0, so that in the management process, more informed decisions can be made when choosing a strategy for the development of regions and enterprises.

In the course of the recent work of Mykola Ivanov, Nataliia Maksyshko, Sergey Ivanov and Nataliia Terentieva [10], 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.

Thus, this study makes it possible to carry out strategic planning for the development of a region or an enterprise, taking into account the intellectual assessment of the economic efficiency of personnel.

2.2 Literature review

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.

The first direction is determined by the level of application of intelligent systems. High information technologies are sweeping the world and replacing the classical methods of managing HR processes. Nowadays, robotic programs are being used that offer employees of the enterprise to pass an express interview or interview using expert systems, which are still in use.

The next direction is intellectual analysis and assessment of the economic efficiency of human resources capable of learning and solving modern problems.

In addition to these areas, one should also take into account modern methods of personnel management, focused on the maximum use of the intellectual capital of employees. This is stimulated by the simultaneous satisfaction of individual needs, desires, and capabilities of employees and their synchronization with the tasks of the enterprise. Modern HR specialists are beginning to more closely monitor the development of employees within the enterprise, which allows flexible management of career growth, which can be adjusted taking into account the proposals of the employees themselves.

Personnel strategy is part of the overall strategy for the development of enterprises and long-term planning, and their business activities. An important role in these plans is played by the assessment of the degree of personnel efficiency as a factor in the renewal and increase in production efficiency in the general economic strategy of the enterprise. Building and managing modern employee data directories requires processing a lot of information. This is due to a wide range of organizational, economic, technical and technological problems solved by the personnel. Therefore, data analysis in HR process management is an urgent task.

However, the problems of rating management have not yet been resolved. The solution to the problem of rating management was the work of Iurii H. Lysenko, Volodymyr L. Petrenko, Oleh I. Bohatov and Volodymyr H. Skobeliev [16]. However, the level of personnel development and their assessment were not taken into account. This work was devoted to the solution of this problem, in which the theoretical aspects of personnel development are studied, in particular the concept, main tasks and directions of personnel development at the enterprise. V. M. Helman, Ye. V. Makazan and A. M. Buriak [6] 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.

Human resource management as a strategic human resource management is to-day considered as going beyond such management tasks as motivation, the level of remuneration. Instead, managers should view human resource management as a process that contributes to the success of the enterprise. Therefore, in the work of Brian E. Becker and Mark A. Huselid [1] 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. In addition, the authors, when solving these problems, considered the issues when the company develops and motivates the development of human capital. They also identified the requirements for businesses that value a skilled workforce and are more profitable than those that do not value this workforce.

The results of the work of scientists Mark A. Huselid [8], Jeffrey Pfeffer and John F. Veiga [20] show that successful enterprises have several characteristics in com-mon: stable job security, high levels of self-government and excellent wages. The most successful businesses manage people as a strategic asset and measure the performance of people in terms of their influence on the manufacturing process. Ibraiz Tarique, Dennis Briscoe and Randall S. Schuler in the work [26] 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 in [25]. 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 Mykola Ivanov, Sergey Ivanov, Nataliia Terentieva, Victoria Maltiz and Julia Kalyuzhnaya [9].

The use of big data in human resource management has been reviewed by Peter O'Donovan, Colm Gallagher, Kevin Leahy, Dominic T. J. O'Sullivan [19] and Alessandra Caggiano [2].

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.

2.3 Formal problem statement

In the Industry 4.0 strategy, human resources are part of the overall development strategy of the state, region and enterprise, and allow them to ensure the current planning of economic activities. An important role in these plans is played by the assessment of the degree of efficiency of human resources as a factor of renewal and increase in production efficiency in the general economic strategy of the region and the enterprise. The creation of modern Big Data about the population, employees and their management requires processing a large amount of information using intelligent assessment using fuzzy logic. This is due to a wide range of tasks solved by the personnel, both organizational, economic and technological. Therefore, fuzzy modeling of big data of human resources in the conditions of Industry 4.0 is an urgent task.

2.4 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.

3 The theory of fuzzy modeling of HR

3.1 Big Data architecture and analysis theory

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 [28], 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).

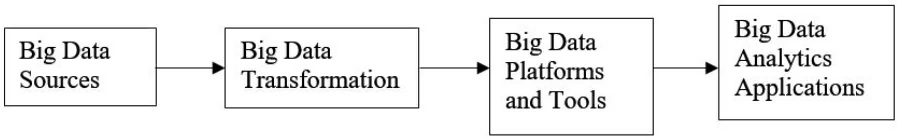


Fig. 1. Architecture Big Data analytics.

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 Data Warehouse has a complex layered architecture called the Layered Scalable Architecture (LSA). In the Big Data system, LSA performs the logical division of data structures into several functional levels. Data is copied (saved) from level to level and transformed at the same time, in order to eventually appear in the form of consistent information. This information will be ready for further analysis.

The key components of Big Data Analytics Applications are Data Mining and OLAP (On-Line Analytical Processing) multivariate data analysis technology. OLAP is a key component of traditional data warehouse organization. OLAP systems are, in one way or another, based on a data storage and organization system.

The development of alternative methods for searching and aggregating information in sparse data hypercubes implies work in a number of directions. Among these areas, one can single out the study of the data model and the formalization of methods for assessing the density of the data hypercube, the study and development of effective methods for accessing information in the data hypercube, the development of an alternative method for aggregating the sparse data hypercube, the study of the possibilities of using various methods of visualization of data hypercubes, and others.

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

$$P = \{p_i = (op1, op2, ep1, ep2, ep3)\},\ i = \overline{1, N}\ , \tag{1}$$

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

$$M_{p_i} = \{m_1,\ m_2, \ldots, m_n\},\ i = \overline{1, n}\ . \tag{2}$$

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

$$M'_{p_i'} = \{m_1',\ m_2', \ldots, m_n'\},\ i = \overline{1, n}\ . \tag{3}$$

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', M' and we will denote it as $H'(D', M')$.

A single set of measurement labels $M_h \subset M$ corresponds to each component of the data hypercube HR $h \in H$. If the HR data cell is empty (does not contain data), then the set of dimensions of the hypercube $H(D, M)$ will be denoted by $V(H)$.

Consider data manipulation operations in a hypercube. We propose a method for managing HR data that includes the following steps.

Stage 1. Data Projection – Multiple Query (MDX).

A subset of the multidimensional data cube $H'(D', M')$ represents a query (Slice).

Stage 2. Building a multidimensional query.

The construction of the query is carried out in order to obtain the necessary subset of the components $H' \subset H$ and remove the values by sequentially fixing the labels. The request is usually an HR dataset.

The label $m_i \in M$ defines the hyperplane of intersection of the data hypercube corresponding to the dimension $d_i \in D$. The set of fixed labels $M' \subseteq M$, thus, defines the set of hyperplanes of sections of the data hypercube, corresponds to the set of fixed dimensions $D' \subseteq D$. The intersection of these hyperplanes determines the set of components of queries of the data hypercube $H'(D', M')$, which is needed by the management level. The essence of the process of extracting data from a hypercube is to construct a slice of the data hypercube $H'(D', M')$ by specifying the sets D', M'.

Tearing out the label $m_i \in M$, corresponding to the measurement $d_i \in D$, we determine what further interests us in the data hypercube. At each next step, the user has access to labels corresponding to the set of unfixed measurements.

Thus, by selecting a label in the dimensions $d_i \in D$ and $d_i \in D'$, we get a query to the HR data hypercube in the form of a table or surface on a MATLAB plot.

The essence of the process of adding money from the hypercube is stored in prompting for the development

Stage 3. Changing the order of presentation (visualization) of measurements.

Changing the order in which measurements are presented is called Rotate. Rotation provides the ability to visualize data in the most comfortable form for their perception. In terms of the data model under consideration, rotation means changing the sequence of fixing marks when building a slice. The result of rotation for a two-dimensional slice (table) will be replacing columns with rows, and rows with columns.

Stage 4. Convolution and detailing.

Convolution and detailing are carried out due to the presence of a hierarchical structure of dimensions. Measurement values (labels) can be grouped into hierarchies consisting of one or more levels. For example, time labels are naturally combined in a hierarchy with levels: year, quarter, month, day. The operations of convolution and detailing do not fundamentally differ from the operation of building a slice of a data hypercube, but they are distinguished to describe the work with aggregated data. The presence of a hierarchical structure of dimensions allows for data aggregation.

Stage 5. Aggregation of tributes.

The number of aggregates stored in the data hypercube along with the primary data depends on the number of labels corresponding to the levels of the hypercube dimension hierarchy, starting with $l = 1$, and can significantly exceed the amount of primary data. The total number of aggregates in the case of two dimensions will be determined by the sum of the values of the areas A_{01}, $A_{02}, \dots, A_{22}$, which are shown in fig. 2.

Filling the HR hypercube with data with an insufficient amount of initial data leads to the formation of empty components. Data hypercubes with many empty cells are sparse.

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.

3.2 Human Resources fuzzy modeling method

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.

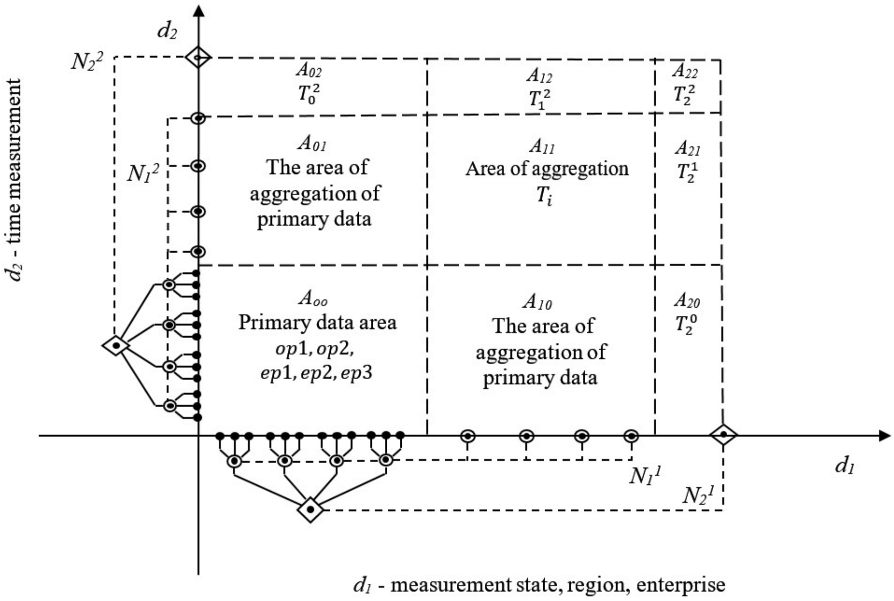


Fig. 2. Aggregation of HR data.

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 proposed, which is presented in fig. 3.

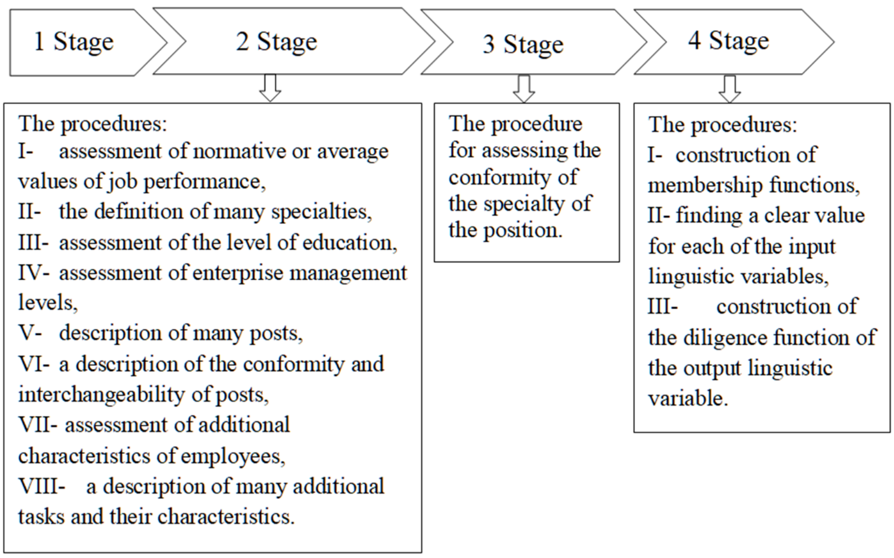


Fig. 3. The method of modelling human resources in fuzzy management.

The proposed method includes four stages.

3.3 Stage 1 Procedures

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

3.4 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 $P = \{pi\ = (op1, op2, op3)\}$, 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 – $P0$:

$$P0 = \{p0_i = (op01, op02)\},\ i = \overline{1, N}, \tag{4}$$

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:

$$SP = \{sp_r\}, r = \overline{1, L^{sp}}, \tag{5}$$

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:

$$UO = \{uo_v = (name, \mu\ \}, v = \overline{1, L^{uo}}, \tag{6}$$

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:

$$UD = \{ud_w = (name, \gamma\ )\}, w = \overline{1, L^{ud}}, \tag{7}$$

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:

$$D = \{d_j = (name, ud, uo, kl, SPD): ud \in UD,\ uo \in UO\}, j = \overline{1, L^d}, \quad (8)$$

$$SPD = \{spd_r = (sp, \beta): sp \in SP,\ 0 \leq \beta \leq 1\}, r = \overline{1, L_j^{spd}}, \quad (9)$$

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^{spd} – number of specialties in the *j*-th position.

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

$$SD = \{sd_f = (d1, d2, \alpha): d1 \in D, d2 \in D\ 0 \leq \alpha \leq 1,\ (d1 = d2) \Rightarrow \alpha = 1\},$$
$$f = \overline{1, L^{sd}}, \quad (10)$$

$$L^{sd} = (L^d)^2, \quad (11)$$

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:

$$A = \{a_i = (ds, ST, OB): ds \in D,\ uo \in UO\}, i = \overline{1, N}, \quad (12)$$

$$ST = \{st_j = (d, kL): d \in D,\ d \in UO\}, j = \overline{1, L_i^{st}}, \quad (13)$$

$$OB = \{ob_w = (sp, uo, god): sp \in SP,\ uo \in UO\}, w = \overline{1, L_i^{ob}}, \quad (14)$$

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:

$$Z = \{z_k = (t0, tk, tk', usz),\ 0 \le \beta \le 1\}, k = \overline{1, M}, \tag{15}$$

where z_k – is the vector of characteristics of the k-th task,

$t0$ and 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:

$$IZ = \{iz_k = (a, z, uv^p, uv): a \in A, z \in Z, 0 \le uv^p \le 200\},\ k = \overline{1, M}, \tag{16}$$

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 $IZ = 0$.

3.5 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:

$$f\beta(xsp, xd) = \begin{cases} d_{j0}.spd_{i0}.\beta, \exists j_0, r_0: (d_{j0} = xd) \wedge (d_{j0}.spd_{r0}.sp = xsp) \\ 0, \qquad \neg\exists j_0, r_0: (d_{j0} = xd) \wedge (d_{j0}.spd_{r0}.sp = xsp) \end{cases}. \tag{17}$$

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

$$f\alpha(xds, xd) = \begin{cases} sd_g\alpha,\ \exists g_0: (sd_{g0}.d1) \wedge (sd_{g0}.d2 = xd) \\ 0,\ \neg\exists g_0: (sd_{g0}.d1) \wedge (sd_{g0}.d2 = xd) \end{cases}. \tag{18}$$

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

$$p_i.op1 = \delta(op11 \cdot op12) \cdot op13, \tag{19}$$

$$op11 = \sum_{w=1}^{L_i^{ob}} f\beta(a_i.ob_w.sp, a_i.ds) \cdot \frac{a_i.ob_w.god}{godT}, \tag{20}$$

$$op12 = \sum_{j=1}^{L_i^{st}} f\delta(a_i.ds.f, a_i.st.d) \cdot \frac{a_i.st_j.kL}{a_i.ds.kL}, \tag{21}$$

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].

3.6 Stage 4 Procedures

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

The following “position”, “level”, “education” can be attributed to numerical linguistic variables of employees, and “conflict”, “level of substitution” to linguistic variables. Numerical linguistic variables and their meanings serve for a qualitative description of a quantitative quantity. The values of linguistic variables are deter-mined by experts.

It should be noted that a linguistic variable, like its original term set, is associated with a specific dimensional scale on which all arithmetic operations are defined.

To assess the characteristics of employees in table 1, linguistic variables and their dimensions are proposed.

The use of the concept of stimulation and destimulation is applied taking into account the influence on the degree of personnel efficiency, namely, stimulation – the effect on the increase and destimulation – on the reduction of the factor.

Therefore, the term set $T_i^n = \{T_i^{n\prime}\}$ is associated with the set $T_i^{n\prime}$, where $T_i^{n\prime} = \langle x, \mu_{T_i^{n\prime}}(x) \,/\, x \in [x_{min}, x_{max}]\rangle$ is a fuzzy number, $i = \overline{1, m}$, m is the number of term sets, n is the number of employees.

To eliminate the influence of changes in the input variables of the metrics and, as a consequence, the correction of term sets, a transition to a normalized function is proposed. Let the previously defined term set T_i be the original one.

The normalized linguistic variable is a mapping on the interval [0; 1]:

$$D_i^n = \{D_i^{n\prime}\},\ D_i^{n\prime} = \langle z, \mu_{D_i^\prime}(z) \mid z \in [0; 1]\rangle, \tag{22}$$

where z – is a fuzzy number corresponding to the term set $D_i^\prime$ on the interval [0; 1], n – is the number of employees.

Table 1. The linguistic variables of employee characteristics.

Term set	The metric and type of exposure	x^1_{min}	x^1_{max}	The term designation
$T_1^1 = \cup T_1^j$, $j = \overline{1,3}$	Performance of duties <position>, stimulation	0	1	Not performed Partially completed Performed
$T_2^1 = \cup T_1^j$, $j = \overline{1,3}$	Job Interchangeability <Interchangeability Level>, discouragement	1	3.0	Low Average High
$T_3^1 = \cup T_1^j$, $j = \overline{1,3}$	Level of education <education>, stimulation	1	3.0	Secondary education Bachelor Master
$T_4^1 = \cup T_1^j$, $j = \overline{1,3}$	Conflict <conflict>, stimulation	0	3,0	Low Average High
$T_5^1 = \cup T_1^j$, $j = \overline{1,3}$	The importance level of the specialty <SpecialtyLevel>, discouragement	0	0,5	Low Average High

These functions allow you to display heterogeneous input variables in a single normalized interval [0; 1], which allows you to reduce errors associated with different quantities and their dimensions. This provides a convenient representation of the values, as well as their interpretation.

4 The method of constructing a model of data mining in HR process management

The structural model of data mining in HR process management is presented in fig. 4.

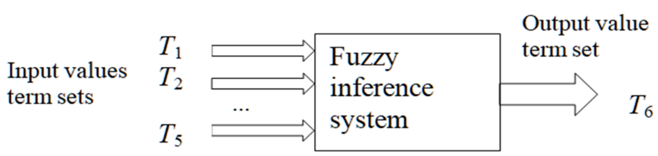


Fig. 4. The structural model of the model of data mining in the management of HR-process.

In the structural model, $T = \{T_i\}$ is a term set, where $i = \overline{1,n},$ n – is the number of sets, each of which is represented by a fuzzy variable with a domain of definition *X*.

The process of modeling fuzzy values is based on a fuzzy inference system, which allows you to convert expert estimates into fuzzy values.

In the fuzzy inference system, the procedure for finding a clear value for each of the input linguistic variables based on defuzzification is applied. Defuzzification in a fuzzy inference system is the process of finding a value for each of the output linguistic variables of the set $W = \{x_1, x_2, .. x_n\}$. The task of defuzzification is to use the results of accumulation of all output linguistic variables. It is necessary to obtain a quantitative value of each of the output variables. Output variables can be used in a fuzzy inference system relative to the input linguistic variable.

Accumulation of fuzzy inference is the process of finding the membership function for each of the output linguistic variables of the set.

The transformation of a fuzzy set into list of values of variables is named as defuzzification.

The defuzzification procedure is performed by a sequence that considers each of the output linguistic variables and the β fuzzy set $T_i = \{T_i^j\}$ related to it. The result of defuzzification for the output linguistic variable is defined as a quantitative value.

The defuzzification process is considered complete when quantitative values are determined for each of the output linguistic variables. For the fuzzy inference system, the Mamdani algorithm was applied [5; 24].

The Mamdani algorithm includes the following steps [17; 3]:

1. the formation of a rule base for fuzzy inference systems [15; 4];
2. fuzzification of input variables;
3. aggregation of conditions in fuzzy rules to find the degree of truth of the conditions of each of the rules of fuzzy logic;
4. accumulation of conclusions of fuzzy production rules;
5. defuzzification of output variables based on the center of gravity method.

An example of a rule looks like this:

```
1. If (Position is Npfd) and (InterchangeabilityLevel is low) and
(Education is secondary) and (Conflict is low) and (SpecialtyLevel is
low) then (StaffEfficiency is NotHardworking) (1)
2. If (Position is Npfd) and (InterchangeabilityLevel is low) and
(Education is bachelor) and (Conflict is low) and (SpecialtyLevel is
low) then (StaffEfficiency is NotHardworking) (1)
3. If (Position is Npfd) and (InterchangeabilityLevel is low) and
(Education is master) and (Conflict is low) and (SpecialtyLevel is low)
then (StaffEfficiency is NotHardworking) (1)
4. If (Position is Pfd) and (InterchangeabilityLevel is high) and
(Education is master) and (Conflict is low) and (SpecialtyLevel is high)
then (StaffEfficiency is Prospective) (1)
```

The values of linguistic variables are determined on an ordinal scale. It should be noted that a linguistic variable, like its original term set, is associated with a specific scale on which all arithmetic operations are defined.

Therefore, the term set $T_i = \{T_i^j\}$ is associated with the set T_i^j, where $T_i^j = \langle x, \mu_{T_i^j}(x) \,/\, x \in [x_{min}, x_{max}] \rangle$, $i = \overline{1, n}$; $j = \overline{1, m}$; n is the number of term sets, m is the number of terms.

A model that satisfies these fuzzy sets is their union:

$$\mu T_i = sup\left(\mu_{T_i^j}(x)\right), \;\; T_i = \cup\, T_i^j. \tag{23}$$

We construct membership functions for the linguistic variable characteristics of employees, presented in table. 1.

The process of converting experts' qualitative assessments into fuzzy quantities consists in mapping the elements of the original term set in the form of constructing membership functions of fuzzy quantities $T_i^j \in T_i$.

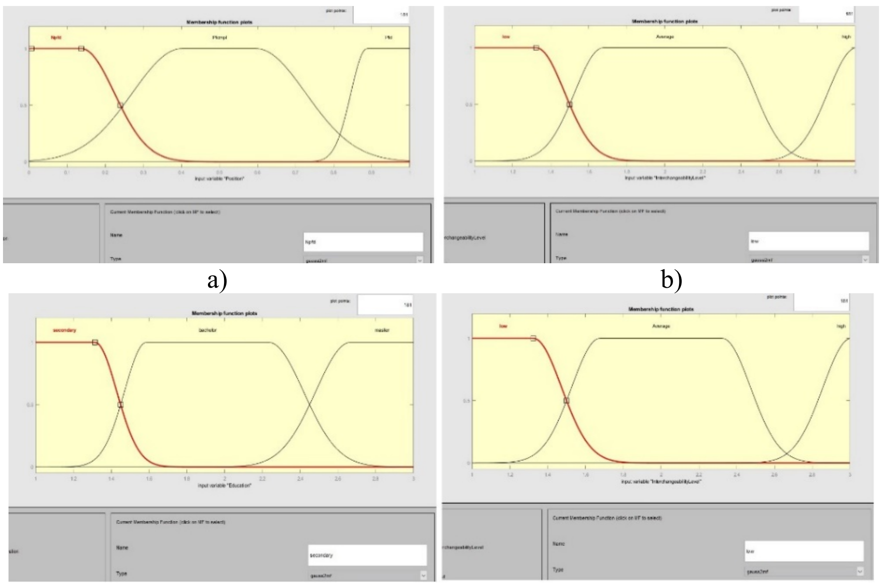
The description of linguistic variables is as follows:

```
⟨Position, {not fulfilled, partially fulfilled, fulfilled}, [0; 1]⟩,
⟨InterchangeabilityLevel, {low, medium, high}, [1; 3]⟩,
⟨Education, {secondary, bachelor, master}, [1; 3]⟩,
⟨Conflict, {low, medium, high}, [1; 3]⟩,
⟨SpecialtyLevel, {secondary, bachelor, master}, [1; 3]⟩.
```

Moreover, the values of the sets are in the range [0; 1] & [1; 3].

5 Experiments and results

The use of the Gaussian function is to use the membership function for modeling to determine fuzzy numbers. It is a form of analytical approximation using functions that include Gaussian functions.



a) b) c) d)

Fig. 5. The membership function of the input linguistic variables: a) "Position", b) "InterchangeabilityLevel", c) "Education", d) "Conflict".

The constructed membership functions of the input linguistic variables are presented in fig. 5.

In the fuzzy inference procedure for managing HR processes, it is necessary to consider the work of employees at all levels of work. The fuzzy inference procedure is implemented in the MATLAB system, which allowed obtaining the following results of assessing the degree of personnel efficiency. To perform the procedure, we built the diligence function of the output linguistic variable “Staff Efficiency”, which is presented in fig. 6.

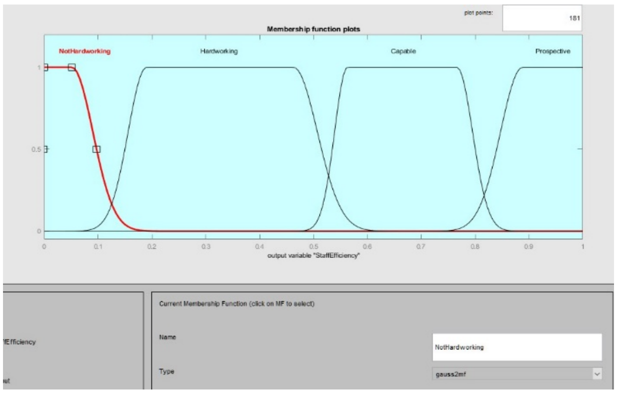


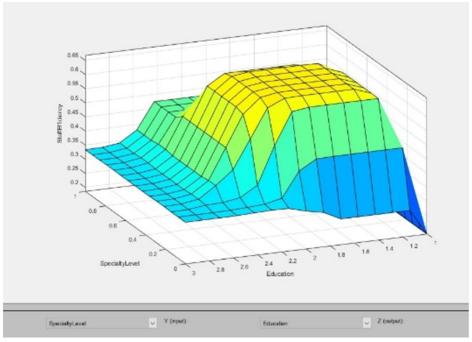
Fig. 6. The membership function of the output linguistic variable “Staff Efficiency”.

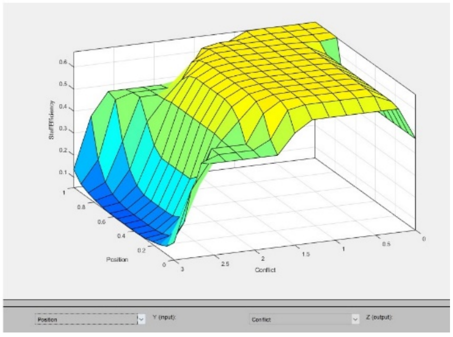
The simulation results of assessing the degree of personnel efficiency, which is presented in fig. 7.

The authors of the article propose a solution to the problem of constructing a data analysis method in human resource management and modeling the assessment of the degree of personnel efficiency based on fuzzy sets.

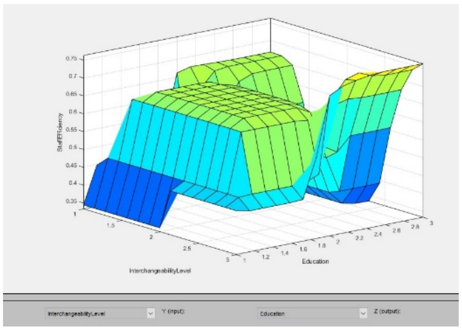
6 Conclusion

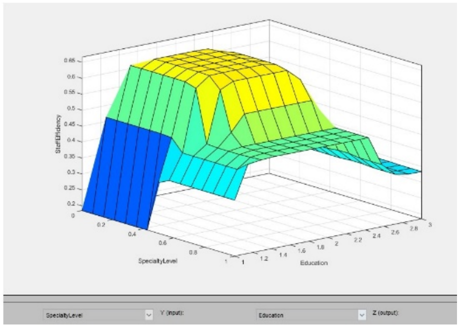
In this article offered the following methods for managing Big HR Data, analyzing and assessing the effectiveness of human resources. Method for analyzing and assessing the effectiveness of HR in four steps is developed. At the first stage, the problem of choosing the analyzed indicators is solved. At the second stage, eight procedures to solve the following tasks are performed. This procedure allows receive assess the standard or average values of labor productivity, determine a variety of specialties, assess the level of education, assess the levels of enterprise management, describe many job responsibilities, 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. In the MATLAB system, a fuzzy inference procedure is implemented, which made it possible to assess the degree of efficiency of human resources.





a) b)





c) d)

Fig. 7. Modeling the assessment of the degree of personnel efficiency: a) the Specialty Level and Education, b) the Position and Conflict, c) the Interchangeability Level and Education, d) the Specialty Level and Conflict.

Prospects for the application of the method and procedures of fuzzy modeling in human resource management lie in the expansion of approaches and the use of models of cloud computing and big data.

References

1. Becker, B.E., Huselid, M.A.: Strategic Human Resources Management: Where Do We Go from Here? Journal of Management **32**(6), 898–925 (2006). doi:10.1177/0149206306293668
2. Caggiano, A.: Cloud-based manufacturing process monitoring for smart diagnosis services. International Journal of Computer Integrated Manufacturing **31**(7), 612–623 (2018). doi:10.1080/0951192X.2018.1425552

3. Dimitrov, V., Borisova L., Nurutdinova, I.: Development and analysis of fuzzy expert data for technological adjustment of a grain harvester header. E3S Web of Conferences **175**, 05027 (2020). doi:10.1051/e3sconf/202017505027
4. Gacto, M., Alcala, R., Herrera, F.: Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. Information Sciences **181**(20), 4340–4360 (2011)
5. Gottwald, S.: Universes of Fuzzy Sets and Axiomatizations of Fuzzy Set Theory. Part I: Model-Based and Axiomatic Approaches. Studia Logica **82**, 211–244 (2006). doi:10.1007/s11225-006-7197-8
6. Helman, V.M., Makazan, Ye.V., Buriak, A.M.: Personnel development as pledge of the success of enterprise. Bulletin of Zaporizhzhia National University. Economic sciences **43**(3), 112–115 (2019). doi:10.26661/2414-0287-2019-3-43-18
7. Hryvko, A., Sytnyk, O., Zhuk, Yu.: Exploring Students' Views on ICT Skills as the Components of Journalists' Professional Competence: EDM Aspect. CEUR-WS.org, online (2020, in press)
8. Huselid, M.: The Impact of Human Resource Management Practices on Turnover and Productivity. The Academy of Management Journal **38**(3), 635–672 (1995). doi:10.2307/256741
9. Ivanov, M., Ivanov, S., Terentieva, N., Maltiz, V., Kalyuzhnaya, J.: Fuzzy modeling in human resource management. E3S Web of Conferences **166**, 13010 (2020). doi:10.1051/e3sconf/202016613010
10. Ivanov, M., Maksyshko, N., Ivanov, S., Terentieva, N.: Intelligent Data Analysis in HR Process Management. CEUR Workshop Proceedings **2608**, 754–768 (2020)
11. Kholoshyn, I.V., Bondarenko, O.V., Hanchuk, O.V., Shmeltser, E.O.: Cloud ArcGIS Online as an innovative tool for developing geoinformation competence with future geography teachers. In: Kiv, A.E., Soloviev, V.N. (eds.) Proceedings of the 6th Workshop on Cloud Technologies in Education (CTE 2018), Kryvyi Rih, Ukraine, December 21, 2018. CEUR Workshop Proceedings **2433**, 403–412. http://ceur-ws.org/Vol-2433/paper27.pdf (2019). Accessed 10 Sep 2019
12. Kiv, A., Semerikov, S., Soloviev, V., Kibalnyk, L., Danylchuk, H., Matviychuk, A.: Experimental Economics and Machine Learning for Prediction of Emergent Economy Dynamics. CEUR Workshop Proceedings **2422**, 1–4 (2019)
13. Kiv, A.E., Soloviev, V.N., Semerikov, S.O.: CTE 2018 – How cloud technologies continues to transform education. CEUR Workshop Proceedings **2433**, 1–19 (2019)
14. Kuyoro, S.O., Ibikunle, F., Oludele, A.: Cloud Computing Security Issues and Challenges. International Journal of Computer Networks (IJCN). **3**, 247-255 (2011)
15. Ludermir, T., Zanchettin, C., Lorena, A.: Advances in intelligent systems. Neurocomputing **127**, 1–3 (2014). doi:10.1016/j.neucom.2013.07.040
16. Lysenko, Iu.H., Petrenko, V.L., Bohatov, O.I., Skobeliev, V.H: Reytingovoye upravleniye ekonomicheskimi sistemami. Yugo-Vostok, Donetsk (1999)
17. Mamdani, E., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies **7**(1), 1–13 (1975)
18. Merzlykin, P.V., Popel, M.V., Shokaliuk, S.V.: Services of SageMathCloud environment and their didactic potential in learning of informatics and mathematical disciplines. In: Semerikov, S.O., Shyshkina, M.P. (eds.) Proceedings of the 5th Workshop on Cloud Technologies in Education (CTE 2017), Kryvyi Rih, Ukraine, April 28, 2017. CEUR Workshop Proceedings **2168**, 13–19. http://ceur-ws.org/Vol-2168/paper3.pdf (2018). Accessed 21 Mar 2019
19. 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). doi:10.1016/j.compind.2019.04.016

20. Pfeffer, J., Veiga, J.F.: Putting People First for Organizational Success. Academy of Management Perspectives **13**(2), 37–48 (1999). doi:10.5465/AME.1999.1899547
21. Schwab, K.: The Fourth Industrial Revolution. Currency, Geneva (2017)
22. Semerikov, S., Chukharev, S., Sakhno, S., Striuk, A., Osadchyi, V., Solovieva, V., Vakaliuk, T., Nechypurenko, P., Bondarenko, O., Danylchuk, H.: Our sustainable coronavirus future. E3S Web of Conferences **166**, 00001 (2020). doi:10.1051/e3sconf/202016600001
23. Semerikov, S.O., Teplytskyi, I.O., Yechkalo, Yu.V., Kiv, A.E.: Computer Simulation of Neural Networks Using Spreadsheets: The Dawn of the Age of Camelot. CEUR Workshop Proceedings **2257**, 122–147 (2018)
24. Siler, W., Buckley, J.: Fuzzy Expert Systems and Fuzzy Reasoning. Wiley Interscience, Birmingham (2005)
25. Stahl, G.K., Björkman, I. (eds.): International human resources management. Handbook of international management research, 2nd edn. Edward Elgar, Cheltenham (2006)
26. Tarique, I., Briscoe, D., Schuler, R.S.: International Human Resource Management: Policies and Practices for Multinational Enterprises, 5th edn. Routledge, New York (2009)
27. 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 **11**, 4–19 (2017). https://doi.org/10.20965/ijat.2017.p0004
28. 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**(1), 2129–2138 (2019). doi:10.1017/dsi.2019.219
29. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Elsevier, San Francisco (2005)