A Comparison of Methods for Identifying the Priority Hierarchy of Influencing Factors

Oleksandr Tymchenko[a,b], Orest Khamula[b], Svitlana Vasiuta[b], Olha Sosnovska[b] and Oksana Mlynko[c]

[a] *University of Warmia and Mazury Olsztyn, Poland*
[b] *Ukrainian Academy of Printing, Lviv, Ukraine*
[c] *Lviv Polytechnic National University, Lviv, Ukraine*

Abstract

This article is a continuation of our research on the identification of factors that affect the visualization of data and the use of infographics. The modern world, and especially its various processes, are changing rapidly. A person needs to perceive more information in a shorter time. Therefore, it is important to study the process of information perception, which should be presented in a clear form and effectively convey the thoughts and ideas of the authors. The large number of factors influencing the process and the complexity of the relationship between them, and consequently the small amount of objective source information do not allow to find the optimal solution.

Based on the selected factors, the experts constructed a dependency graph of relationships between them and performed calculations, which were used to build and determine the priority of the factors and the interdependence between them. The analytic hierarchy process and ranking method were used in this work to calculate and determine the priority of factors influencing the compositional design of infographics with elements of visual communication. Six hierarchy levels of the multilevel model of weight values of influencing factors were obtained using the ranking method. In contrast, by using the analytic hierarchy process, only four levels were obtained. After optimizing the model of the hierarchy of factors influencing the design of infographics with elements of visual communication, five levels were obtained. These calculations showed that, in this case, the ranking method was more significant for the relationships and influences between the factors.

Keywords

Infographics, Data Visualization, Influencing Factors, Dependency Graph, Model, the Analytic Hierarchy Process, Ranking Method

1. Introduction

The importance and significance of information in the modern world should not be underestimated, especially during information warfare. The timeliness of its reception and instant processing allows reacting faster or taking appropriate actions. Of course, the information is diverse and has a different purpose. It can be narrative, present some material in numerical values, or in the form of specific comparisons. The more extensive information flow with certain numerical values, the more difficult it becomes to comprehend at once. It takes some time to perceive and understand it. In such cases, infographics allow visualizing and presenting the information in a more understandable format for perception.

IntelITSIS'2022: 3nd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 23–25, 2022, Khmelnytskyi, Ukraine
EMAIL: olexandr.tymchenko@uwm.edu.pl (O. Tymchenko); khamula@gmail.com (O. Khamula); svitlanavasyta@gmail.com (S. Vasiuta); olhakh@gmail.com (O. Sosnovska); oxanamlunko@gmail.com (O. Mlynko);
ORCID: 0000-0001-6315-9375 (O. Tymchenko); 0000-0003-0926-9156 (O. Khamula); 0000-0003-0078-9740 (S. Vasiuta); 0000-0001-5413-2517 (O. Sosnovska); 0000-0001-9878-6846 (O. Mlynko)

Research on what infographics and their varieties are currently are in some way represented. Thus, in [1], the author provides information about the importance of infographics and their presentation. In our previous work [2], we also provided specific options for information design and areas where it can be used. A meaningful study [3] is the analysis of four key elements of information visualization — text, images, data, and interaction. This study reveals the impact between these elements and the possibility of using the results obtained in learning using the latest technologies. Publications related to infographics can be divided into three groups: the first is devoted to the processes related to human perception and understanding of information [4, 5]; the next group of publications is devoted to methods of reproduction and implementation of infographics [6, 7]; the last group of works considers the question of what data shall be used to create infographics [8, 9]. Of course, this is not a complete list of materials that reveal the nature and purpose of the use of infographics.

All reviewed materials of the conducted research help to understand the need to address the issue of information perception and its presentation in a clear and comprehensible format. Based on the number of publications, it can be noted that this issue is quite actual. In turn, it can be argued that infographics and data visualization is a rather complex process that is influenced by a number of factors, both external and internal. This issue, as literature research has shown, is currently insufficiently explored.

2. Methods for identifying the priority hierarchy of influencing factors, Literature review

To study the influence processes of external and internal factors on the creation and use of infographics, namely the identification of essential parameters in the choice of infographics and data visualization, it was proposed to use methods used in the study of processes. The parameters of these processes cannot be represented in numbers. There are many methods. In this paper, two widely used methods are analyzed — the hierarchy analysis and the method of ranking factors. These two methods have proven themselves in practice.

A. The analytic hierarchy process

The analytic hierarchy process was proposed by the American mathematician T. Saaty. This method is based on pairwise comparisons of factors that were identified with a high priority in a particular process. Based on this, the priority of decisions is determined. T. Saaty proposed to use the so-called pairwise comparison scale to present the results of estimates in quantitative terms. By using this scale, participants conduct a comparison of the relationship among the factors, the influence of one on another. At the same time, we are not interested in the absence of physical or objective units of measurement. The main advantage of this method is that it is dimensionless. It allows ignoring the question of equating the values of criteria or factors to the same units of measurement. The legitimacy of the use of this scale has been repeatedly proven theoretically as well as practically.

Among the research devoted to the use of the analytic hierarchy process are a number of meaningful works. Thus, in [10], a comprehensive study was conducted to analyze the reliability of complex systems with common causes of failure and mixed uncertainty. The importance and sensitivity of different components types and their impact on system reliability were revealed. In turn, the work [11] was also devoted to calculating the system's reliability. The results showed that the proposed method allows to effectively describe the change in system behavior and obtain its reliability by calculating the proposed model. An interesting study that uses hierarchy analysis is research [12], which presents an improved method of early risk prevention to identify food safety. This study shows that the proposed method allows to scientifically and reasonably determine the level of information about the level of risk and provides risk management to effectively reduce risk losses of the country through appropriate quality control departments.

B. The method of ranking factors

The need to anticipate certain situations or processes, as well as to determine the prospects of specific decisions, prompted the development and improvement of forecasting. The vast majority of processes are related to the lack or limitation of initial information required to make a specific forecast. This situation has led to the development and improvement of forecasting methods based on expert assessments. Expert assessment methods (heuristic methods) have developed especially in

recent decades. The method of collective expert assessment attracts special attention. This method is based on the hypothesis that the experts selected for the survey have the appropriate knowledge and ability to choose the most optimal factors (parameters) from the alternatives. This survey is conducted in the form of questionnaires, where experts provide the answers. All these surveys are conducted anonymously, rejecting the collective component in solving a particular process.

In these questionnaires, experts are given the opportunity to evaluate the relative importance of certain factors (parameters) on a 100-point scale. Zero is given to factors that, according to experts, do not affect the process, respectively, 100 points are given to the most important factor. Some factors may have the same number of points. After agreeing on the importance of factors, the next step is to conduct specific calculations, which will be described in the practical part of our article. Multicriteria analysis usually offers a quantitative approach to facilitate decision-making by ranking alternatives. However, when evaluating the importance of criteria and the adequacy of each alternative to each criterion, uncertainty may arise due to two factors. First, expert responses are usually expressed in linguistic terms that do not have a unique quantitative assessment. Second, there may be uncertainty about the answer. Most multicriteria procedures combine fuzzy numbers and linguistic scales to deal with the first factor but underestimate confidence issues.

The studies on the use of the ranking method are also quite noteworthy. The results obtained in [13] show that this ranking method helps decision-makers choose the most reliable alternative since it is possible to eliminate significant differences in the rating with and without uncertainty. The authors state that this method shows great accuracy in modeling uncertain opinions and providing more useful and additional information to better facilitate decision-making. In [14, 15], it is shown that new ranking methods are developed and created for the convenience of calculation and its flexibility. It is indicated that the choice of method has a significant impact on the rating of influencing factors, identifying procedures that offer similar results or differ significantly in terms of the recommended procedure. Another study that indicates the importance of using the ranking method in research of this type is [16]. It is dedicated to the ranking of goods based on online reviews to support consumer decisions to buy online. The study considers the process of combining information for product rating. It consists of three stages: the selection of product characteristics, mood analysis, and product rating.

Based on the review, each method has its aspects and advantages. Therefore, our study compares two methods when deciding which one is more convenient. The calculations are carried out in the study of the data visualization process in infographics to identify the priority hierarchy of influencing factors.

3. Prioritization of factors influencing the process of data visualization in infographics

Successful visualization or presentation of data in infographics is a rather complex process, not yet fully studied. It is known that the person receives 90% of information through sight and 10% through other senses. It raises the issue of creating a clear infographic that will make it easier to understand and comprehend. Currently, it is one of the best visual tools used to attract and retain attention, and an effective presentation of various data. By using infographics, the relevant content can be properly dosed — by combining text with graphics.

A certain array of data can be visualized differently. It can be represented as tables with numerical values, demographic information, web statistics, and many other forms. However, there is an issue regarding the presentation of this information qualitatively since it must be displayed as accurately and clearly as possible. Based on the literature analysis, there is a limited amount of data on when and what type of data visualization should be proposed and what factors affect it.

Students, teachers, and our graduates who work in the field of advertising and IT were surveyed to determine the factors that affect the visualization of data in the infographics. From the study [2], it was obtained a list of the most significant factors influencing the presentation of data visualization in infographics. For better clarity, each factor will be assigned a number:

h_1 – text (T);
h_2 – numerical data (ND);
h_3 – graphs and charts (GC);

h_4 – flowcharts (FC);
h_5 – image (IM);
h_6 – icons (IC).

After receiving the given information on the quantity, weight, and influence of one factor on another, it is necessary to construct the initial graph (see Figure 1) for our further research.

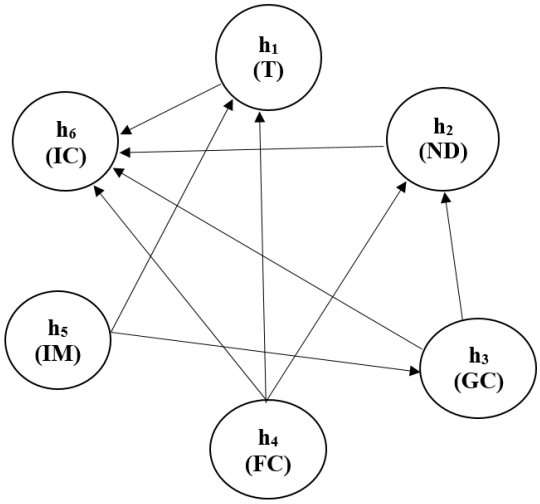


Figure 1: The initial graph of relationships among the factors of infographic design with elements of visual communication [2]

The obtained initial graph is the basis for our further calculations considering the use of the proposed methods to determine the priority of factors influencing the process of data visualization in infographics.

3.1. The Analytic Hierarchy Process

The method is based on the construction of binary matrices of dependencies and reachability among factors and the subsequent definition of a certain level of the hierarchy of priority action of factors [17]. Since this method and the optimization of the model of factors influencing the process of presenting data visualization in the infographics were thoroughly described in [2], our study will focus only on the results obtained using this method.

According to the calculations, the following optimized hierarchical model of priorities of factors influencing data visualization in infographics is obtained (see Figure 1).

According to the obtained model, the factors influencing the process of presenting data visualization in infographics are divided into 5 levels. Icons are the most important factor, then numerical data, graphs and charts, and text. There are two factors of the least level of importance: flowcharts and graphs. From the obtained results, it can be noted that it is necessary to choose icons when starting the process of information visualization in infographics. It, in turn, will help with the identification and unification of information blocks, drawings, and diagrams.

3.2. The Method of Ranking Factors

The initial graph (see Figure 1) is the basis for the implementation of the ranking method, as well as for the analytic hierarchy process. The partial graphical models (hierarchy tree structures) are built based on the obtained initial graph of relationships among influencing factors (see Figure 1 and Figure 1). The models will reflect the hierarchy of influences or dependencies between these factors. Each partial graphical model will become an information input base of numerical parameters for the possibility of obtaining quantitative parameters of these factors and establishing their ranks. When constructing hierarchical trees, both direct and indirect (have their influence due to another factor)

types of influences (see Figure 1) should be considered. It is also necessary to construct direct and indirect dependencies (see Figure 1).

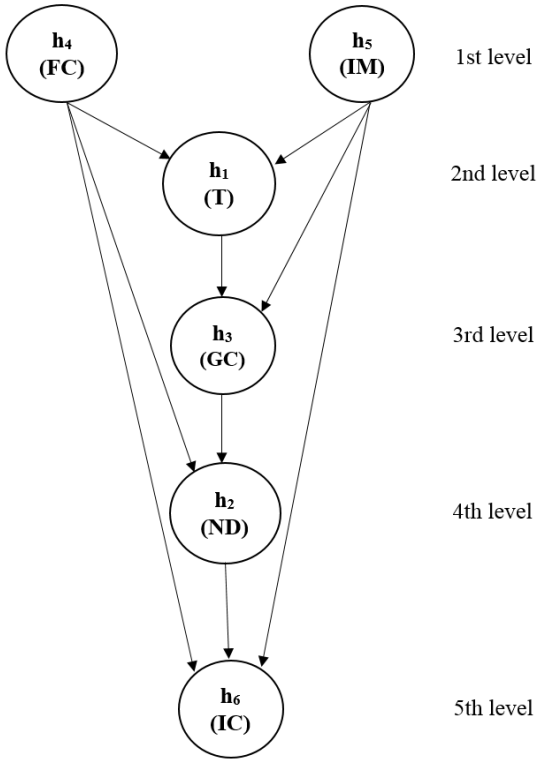


Figure 2: An optimized model of factors of infographics compositional design with elements of visual communication

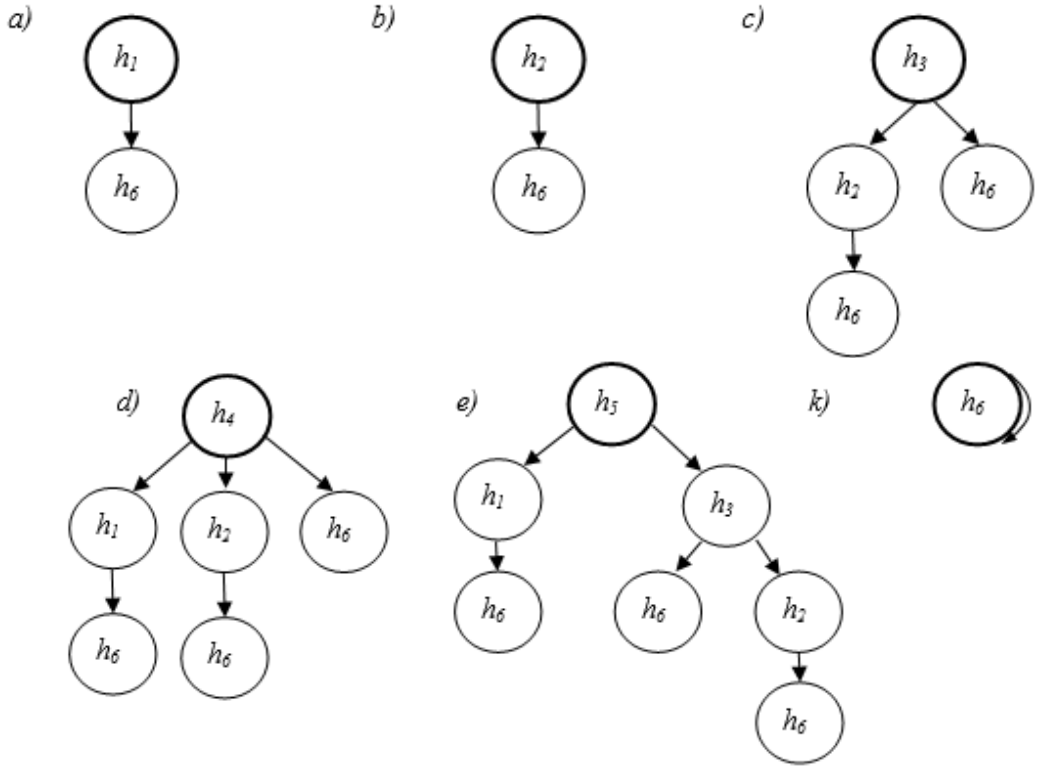


Figure 3: A graph of hierarchical multilevel influences among data visualization factors in infographics

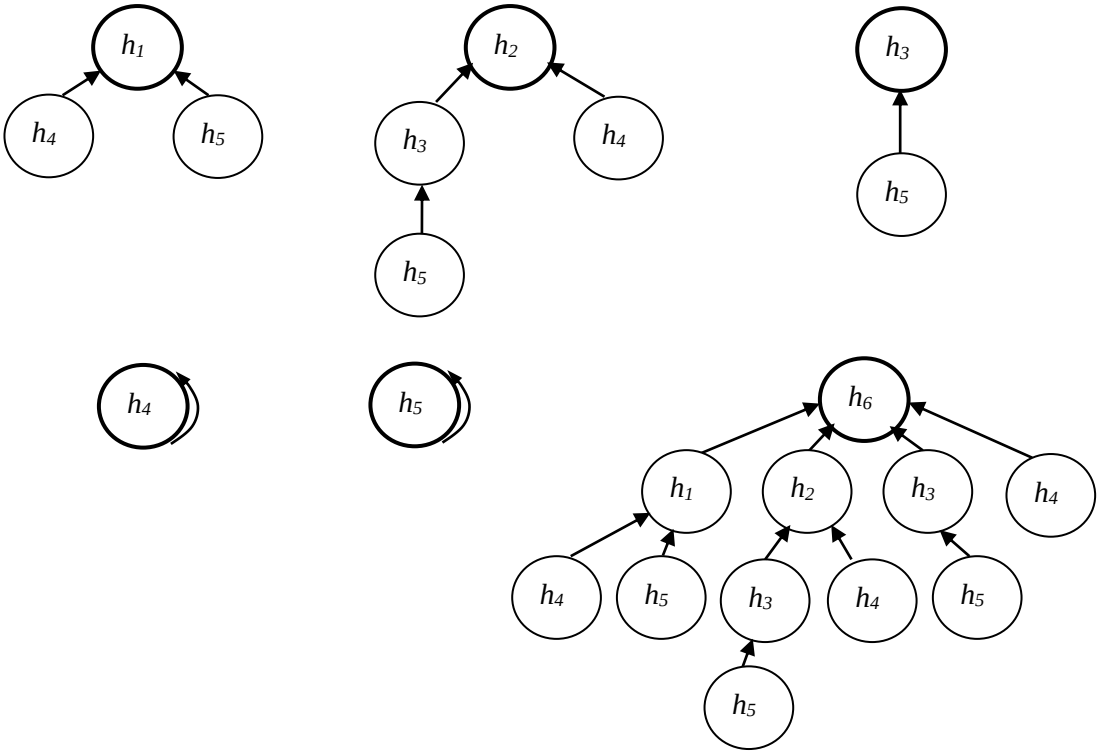
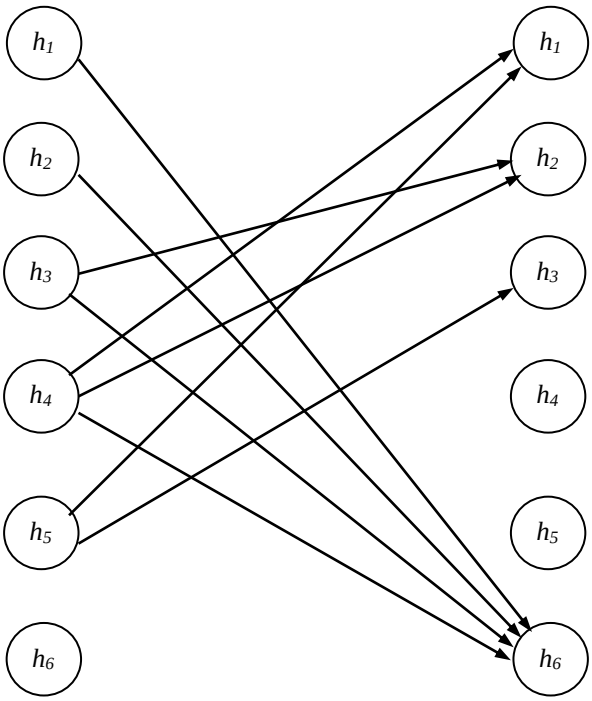


Figure 4: A graph of hierarchical multilevel dependencies among data visualization factors in infographics

A modified scheme of relationships among factors in Table 1 is built based on the obtained graph of relationships among factors influencing the data visualization process in the infographics (see Figure 1). The table shows the factor number and the direction of direct influence of each factor with the way of dependence on other factors.

Table 1

A modified scheme of representation of the relationships among factors

Factor number	Direction of influence	Ways of dependence
1		
2		
3		
4		
5		
6		

First, it is necessary to calculate the total weight values of direct and indirect influences of factors and their integral dependences on other factors. To do this, let's introduce the following notation: g_{ij} – is the number of influences (i = 1 – direct, i = 2 – indirect) and dependencies (i = 3 – direct; i = 4 – indirect) for j^{th} factor (j= 1, ..., n); assuming that w_i – is the weight of the i^{th} type, we obtain: w_1 = 10, w_2 = 5, w_3 = -10, w_4 = -5 respectively, conventional units.

The total weights, in turn, are denoted by K_{ij}. According to the graph theory, by fulfilling the requirements of usage, we obtain the following calculation formulas:

$$K_{ij} = g_{ij} w_i \ (i=1, 2, 3, 4; j= 1, ..., n) \quad (1)$$

where n – is the factor number.

The obtained initial graph of relationships between factors (fig. 1) can be considered as a certain semantic network, based on formula (1) we obtain the following equality:

$$K_{ij} = \sum_{i=1}^{4} \sum_{j=1}^{6} g_{ij} w_i. \quad (2)$$

For our calculations, the following is accepted - in case of the absence of a certain factor of one of the listed relationships types, the corresponding value of g_{ij} in the equality (2) will take the value of zero. By using this formula, we can calculate the weight values of the ranking of factors, taking into account the different types of relationships between them (Table 1).

To build the table of a modified scheme of relationships among factors, in the column "directions of influence," let's choose direct influences for each factor. The number of direct influences is fixed by the coefficient g_{1k}.

The column "dependency paths" allows to obtain the coefficients g_{3k}, and the combined use of indirect influences of the factor or its dependencies allows to obtain the coefficients g_{2k} and g_{4k}.

It should also be noted that the values of g_{3k} and g_{4k} are taken as < 0 according to the given initial conditions $w_3 < 0$, $w_4 < 0$. Accordingly, to reduce the total weight values, formula 2 takes the following form:

$$K_{Fj} = \sum_{i=1}^{4} \sum_{j=1}^{6} g_{ij} w_i + max|K_{3j}| + max|K_{4j}|. \quad (3)$$

Based on this formula, to establish the ranks of experimentally determined factors, let's fill in Table 1.

Table 2

Factor ranking calculations data

Factor number j	g_{1j}	g_{2j}	g_{3j}	g_{4j}	K_{1j}	K_{2j}	K_{3j}	K_{4j}	K_{Fj}	Factor rank (level)	Priority level
1	1	0	2	0	10	0	-20	0	45	3	4
2	1	0	2	1	10	0	-20	-5	40	2	5
3	2	1	1	0	20	5	-10	0	70	4	3
4	3	1	0	0	30	5	0	0	90	6	1
5	2	2	0	0	20	10	0	0	85	5	2
6	0	0	4	3	0	0	-40	-15	0	1	6

From the obtained results, max $|K_{3j}|$ = 40, max $|K_{4j}|$ = 15 (see Table 1), let's sum these two values and add them to the sum of the values K_{1j}, K_{2j}, K_{3j}, K_{4j}.

After performing this calculation, we obtain the value of K_{Fj} – the basis for establishing the rank of factors, and accordingly the level of priority of the impact on the process of data visualization in the infographics.

After obtaining the values shown in Table 1, the next step is to build a multilevel model of the weight values of the influencing factors on the process of data visualization in the infographics. (see Figure 1):

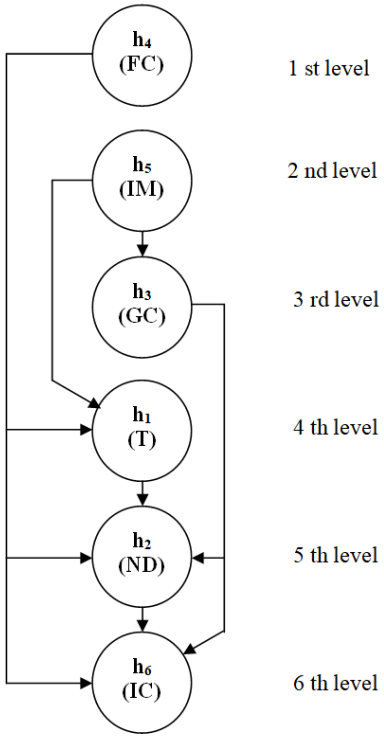


Figure 5: A multilevel model of weight values of influencing factors on the process of data visualization in infographics

4. The results of a comparison of methods for identifying the priority of influencing factors

As a result of our calculations, we obtained the weight values of the factors influencing the process of data visualization in the infographics by the method of ranking factors expressed by numerical values. According to the obtained results, the distribution of influencing factors at six levels of importance was obtained.

As in the previous case, the most important factor was the factor responsible for the choice of icons. Based on this, it is necessary to choose icons when starting the process of information visualization in infographics. It, in turn, will help with the identification and unification of information blocks, drawings, and diagrams.

Numerical data, respectively, are as subsequent in terms of importance. It should also be noted that by using this method, the factors of graphs and charts, and text have changed places by weight. Based on the calculations, according to the ranking method, the text factor has become more important. Flowcharts are on the least level of importance.

This method, in our opinion, and based on the results obtained, allows identifying hidden relationships between these factors.

5. Conclusions

During the practical part of our research, namely identifying the importance of influencing factors in the process of data visualization in the infographics, two widely used methods were used in this work, such as the analytic hierarchy process and the ranking method.

With the help of expert surveys, the influencing factors on the data visualization in the infographics were identified.

Among which the most important are highlighted, namely: h_1 – text (T); h_2 – numerical data (ND); h_3 – graphs and charts (GC); h_4 – flowcharts (FC); h_5 – image (IM); h_6 – icons (IC). The model of influencing factors on the data visualization in the infographics is synthesized by constructing an initial graph of relationships among the selected factors.

Based on the calculations, using two methods, the levels of priority of the influencing factors on the process of data visualization in the infographics are established. The ranking method showed the distribution of factors at 6 priority levels, and the analytic hierarchy process showed the distribution at 5 priority levels.

The study results are presented in the form of multilevel models of weight values of influencing factors. As a result of comparing the obtained calculations, it can be stated that in determining the importance of influencing factors on the data visualization in the infographics, the analytic hierarchy process is inexpedient in use since it does not take into account indirect influences and dependencies, which give preference to one factor over another, and allows the placement of factors on the same level of priority.

By obtaining the study results using two methods, it can be noted that the main advantage of infographics — is to turn uninteresting, incomprehensible information into a graphically structured model, with which every person can understand the content and basic idea of infographics.

Clearly, the artistic aspect of infographics has to be mentioned as well. Studies show that important aesthetic aspects of infographics are using a single font style, the compatible color combination, and the skillful arrangement of infographics elements.

Using well-known learning icons, the effect of memorization and motivation to learn specific information on a specific topic, presented in the style of infographics can be enhanced.

Therefore, following these rules when preparing infographics will help ensure a high-quality product to achieve the goal.

6. Acknowledgments

The authors are appreciative to the large audience of students and teachers who contributed to the research and identification of important factors influencing the choice and creation of infographics, as well as to colleagues for their valuable advice, which allowed improving the quality and enhancing the article`s materials.

7. References

[1] O. V. Gudyma. Infohrafika: navchalnyi posibnyk, Chernivtsi: Chernivetskyi nat. un-t., 2017, 107 p.

[2] O. Tymchenko, N. Kunanets, O. Khamula, S. Vasiuta, O. Sosnovska. Synthesis and research of a model of factors of infographics compositional design with elements of visual communication. Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security with CEUR-WS Khmelnytskyi, Ukraine, March 24–26 (2021) 303-322.

[3] M. Sorapure. Text, Image, Data, Interaction: Understanding Information Visualization. Computers and Composition. 54 (2019) 10251. doi: 10.1016/j.compcom.2019.102519.

[4] I. Tollis, K. Kakoulis. Algorithms for visualizing phylogenetic networks. Theoretical Computer Science, 835 (2020) 31-43. doi:10.1016/j.tcs.2020.05.047.

[5] M. Wang, B. Yuan, P. A. Kirschner, A. W. Kushniruk, J. Peng. Reflective learning with complex problems in a visualization-based learning environment with expert support. Computers in Human Behavior. 87 (2018) 406-415. doi:10.1016/j.chb.2018.01.025.

[6] SH. Cheing, Si Yain-Whar, R. K. Wong. Online force-directed algorithms for visualization of dynamic graphs. Information Sciences, 556 (2021) 223-255. doi:10.1016/j.ins.2020.12.069.

[7] A. Moraes, B. Rodrigues, G. Diniz, J. Barbosa, H. Côrtes, V. Lopes, S. Diniz, J. Barbosa. What questions reveal about novices' attempts to make sense of data visualizations: Patterns and misconceptions. Computers & Graphics, 94 (2021) 32-42. doi:10.1016/j.cag.2020.09.015.
[8] AM. Rodrigues, GD. Barbosa, H. Lopes, SD. Barbosa. Comparing the effectiveness of visualizations of different data distributions. In: Proceedings of the 32nd SIBGRAPI conference on graphics, patterns and images (SIBGRAPI), (2019) 84–91. doi:10.1109/SIBGRAPI.2019.00020.
[9] B. Pinaud, J. Vallet, G. Melançon. On visualization techniques comparison for large social networks overview: A user experiment. Visual Informatics. 4 4 23-34. doi:10.1016/j.visinf.2020.09.005.
[10] J. Mi, N. Lu, Y.-F. Li, H.-Z. Huang, L. Bai. An evidential network-based hierarchical method for system reliability analysis with common cause failures and mixed uncertainties. Reliability Engineering & System Safety, 220, 108295. doi:10.1016/j.ress.2021.108295.
[11] Y. Chen, Y. Yi Li, R. Kang, M. Ali. Reliability analysis of PMS with failure mechanism accumulation rules and a hierarchical method. Reliability Engineering & System Safety, 197 (2020) 106774. doi:10.1016/j.ress.2019.106774.
[12] B. Ma, Y. Han, S. Cui, Z. Geng, H. Li, C. Chu. Risk early warning and control of food safety based on an improved analytic hierarchy process integrating quality control analysis method. Food Control, 108 (2020) 106824. doi:10.1016/j.foodcont.2019.106824.
[13] M. Juanpera, B. Domenech, L. Ferrer-Martí, A. García-Villoria, R. Pastor. Methodology for integrated multicriteria decision-making with uncertainty: Extending the compromise ranking method for uncertain evaluation of alternatives. Fuzzy Sets and Systems. Available online 17 August 2021. doi:10.1016/j.fss.2021.08.008.
[14] O. Apaydın, Z. Aladağ. Ranking the evaluation criteria of Hi-Fi audio systems and constricted information space: A novel method for determining the DEMATEL threshold value. Applied Acoustics. 190 (2022) 108584. doi:10.1016/j.apacoust.2021.108584.
[15] A. Labijak-Kowalska, M. Kadziński. Experimental comparison of results provided by ranking methods in Data Envelopment Analysis. Expert Systems with Applications, 173 (2021) 114739. doi:10.1016/j.eswa.2021.114739.
[16] Z.-P. Fan, G.-M. Li, Y. Liu. Processes and methods of information fusion for ranking products based on online reviews: An overview. Information Fusion, 60 87-97. doi:10.1016/j.inffus.2020.02.007.
[17] O. Tymchenko, S. Vasiuta, O. Khamula, A. Konyukhov, O. Sosnovska, M. Dudzik. Synthesis of Factors Model of Data Visualization in the Infographics. 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications Science and Technology PIC S&T'2019: Conference Proceedings. Volume 2. (October 8-11, 2019, Kyiv, Ukraine). Kyiv, (2019) 451-454. doi:10.1109/PICST47496.2019.9061304.