Visualization in a Knowledge Transfer Process

Janusz Opiła
AGH University of Science and Technology, Cracow, Poland

Abstract

Visualization is not only an important part of experimental data analysis process but also an efficient tool for transfer of acquired knowledge in multiple areas including, but not restricted to, ICT, general scientific discussion, innovation, data mining, decision support systems as well as education and learning. Consequently search for new and perfection of present visualization techniques is a necessity. In the paper, there are discussed selected visualization designs, doable using OpenSource tools such as ScPovPlot3D library, graph-drawing program „graphviz" or charting application „gnuplot". Presented ideas are discussed using two examples. The first one is a theoretical tool for consumer sentiment analysis, forecasting and what-if analysis focused on tourism. The second one is usage of graphing utility "graphviz" for visual analysis and revealing of hidden organisational structures.

Keywords: Visualization, Knowledge Transfer, Simulation, Sentiment Forecasting, graphviz.
JEL classification: C01, C15, C44, C51, C53, D81, D83, D87, D91, L22, L25, Z32

Acknowledgments: The work was supported by AGH University of Science and Technology under grant: 11/11.200.327, job 5.

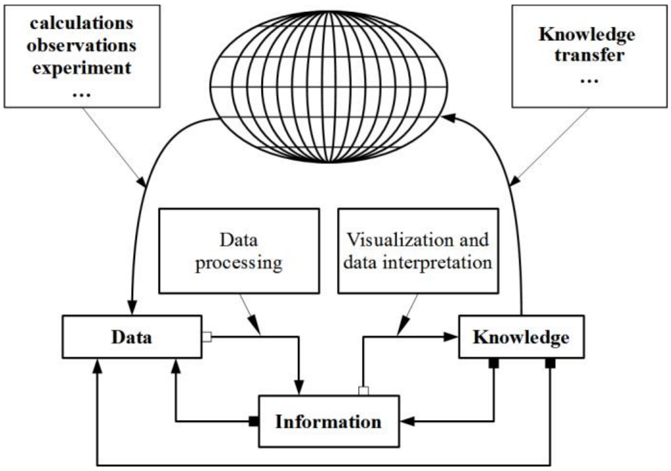
Introduction

Innovations are the important product of the information society - while they accumulate, they gradually form the basis for a knowledge-based economy. This applies not only to so-called high technologies, such as aeronautics, biomedical engineering or the microelectronics, but also to the use of biological natural resources, e.g. for fishing or agricultural and forest economy but includes services like tourism as well. Acquiring knowledge is a complex, multi-stage process, and its inherent part is the transfer of knowledge between individuals. Without knowledge sharing, collected knowledge would be a valuable, though useless, collection of personal notes. Acquiring and transferring knowledge is a complex and continuous process involving many techniques of expression, including oral communication, writing as well as numerous forms of visualization (Opiła, 2017b).

The process of acquiring and transferring knowledge

Most of the cognition results from the observation of nature, which is rather passive in descriptive sciences (geography, astronomy, history, mathematics, economy) or active, in experimental ones (physics, biology, chemistry, engineering). With a relatively high level of generality, it can be traced with help of Fig. 1.

Figure 1
Knowledge Transfer Scheme



Source: Author's illustration, (Opiła, 2017b) reworked

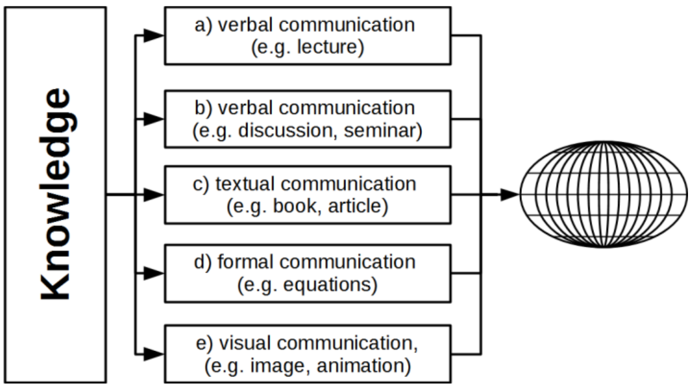
The very first stage is usually the reception of the observable environment which leads to the formation of data sets - consisting of remarks, notices, photographs, maps, measurements. At this stage, raw data have no meaning thus the next step is data processing. Besides a simple classification of data, advanced techniques might be employed such as multidimensional statistical analysis, modelling and visual data analysis (Opiła, 2005, Opiła, 2015, Opiła, 2016a, Opiła, 2016b, Kozlíková et al., 2017). „*When data are processed, interpreted, organized, structured or presented so as to make them meaningful or useful, they are called information*" (Diffen, 2018). Acquired information is basis for extension of knowledge, but on the other hand it can be used for deeper understanding and subsequent re-analyze of the data-set or to obtain additional data based on more precise or complementary measurements, e.g. elevation scanning using LIDAR after geodetic measurements on a triangulation grid. Again, acquired knowledge can be used as a guide to the re-analysis of information or impulse to collect additional data - this time being focused on testing of research hypotheses or simply observing the dynamics of the evolution of the monitored system. However, the acquired knowledge is still hermetically sealed inside the research team. For this knowledge to become useful, one more step is necessary - knowledge transfer, firstly between specialists, secondly between research institutions, thirdly from professors to their students and finally to general public.

The discovery of specific patterns of bird migrations (Opiła, 2016c), will remain a dead record in library until it is transferred to dedicated agencies, such as air traffic surveillance or agricultural agencies suggesting farmers the dates of spraying so as not to jeopardize migratory birds. Determining the reasons for the increase in the level of methane, a potent greenhouse gas, will not help until they are successfully passed to decision makers. Knowledge of the dynamics of tourist traffic at the level of a scientific institute will not help apartment owners to react adequately to disturbances caused by, for example, unfavourable weather until they are properly notified about possible threats. Only then, they can counteract the decline in profits through properly constructed promotions dedicated for selected groups (e.g. young couples) or well calculated rental price cuts.

Data visualization and the knowledge transfer

Knowledge transfer is a complex process, prone to distortions and requiring the sender and recipient to share the same conceptual apparatus. Discussion considering the influence of pollutants, such as bisphenols (BPA), on the body's functions, including neurohormone system, requires that both the sender and the recipient have knowledge of bisphenols chemistry (Gosztyła et al., 2011), vectors of immission, their biological activity and retention in natural environment, patterns of introduction into the organism - with food, inhaled air or maybe transdermally - distribution patterns, cellular metabolism and ultimately excretory routes.

Figure 2
Knowledge Transfer - Selected Ways



Source: Author's illustration (Opiła, 2017b), reworked

Thus, the traditional textual transfer of knowledge often becomes ineffective, because it is impossible to efficiently present the spatial relationships in a chemical molecule, discuss the distribution of electron density or indicate the energy levels of the molecules responsible for biochemical reactions. However, also at microscopic levels, the limitation to verbal communication may hinder the perception of the obtained research results. The conclusion can be drawn that valuable work deprived of visualization becomes hermetic even for specialists in related fields, not to mention representatives of other scientific disciplines. Although this is not a work fault, as the insightful reader can find suitable schemes in the literature cited, it can introduce a significant delay in understanding the material covered. A graphical diagram of selected knowledge transfer paths is shown in Fig. 2 (Opiła, 2017b). In general, it was decided to distinguish five paths of knowledge transfer:

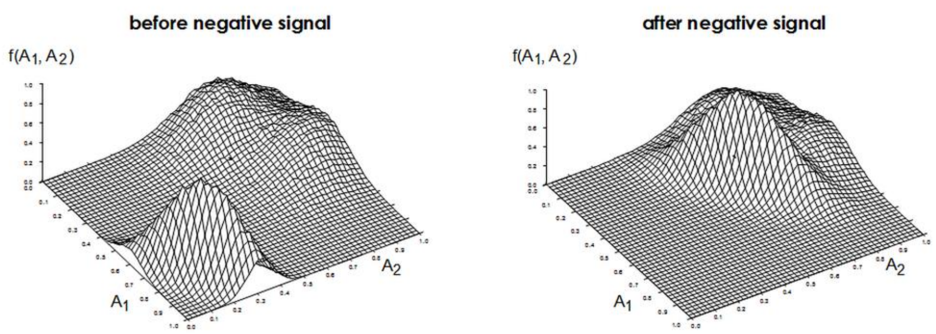
- verbal - non-interactive (e.g. lecture, presentation),
- verbal - interactive (e.g. seminar, panel),
- static – mainly printed text (books, textbooks, e-books),
- highly formalized (e.g. mathematical formulas or music notation),
- visual message (static graphics, 2D or 3D, Virtual Reality or animation).

Each of these techniques has its advantages and disadvantages and the field where it can or should be used. Although the lecture allows for an orderly transmission of content, the possibilities of asking a lecturer are limited and losing a thread often causes a misunderstanding of the whole issue. Though the seminar allows for discussion, it is effective only in a small group and there is always a risk of discussion about personal ambitions. The printed text is usually carefully thought out, orderly, available to many recipients and allows re-analysis of the content to a full understanding of the argument, on the other hand, the rate at which knowledge increase means that the finished text is outdated and its specificity makes it impossible to ask the author for details of his deduction paths. A strictly formalized message is admittedly precise but its proper interpretation requires training and knowledge of certain conventions specific for the given specialty. In this context, the visual message seems to be the most attractive and accessible to a mass audience. However, there are also strong constraints here. The specificity of a scientific discipline may require visualization in compliance with strictly defined rules, understandable only for the educated recipient. The structural formula of the bisphenol molecule (Kuźmińska-Surowaniec et al., 2016) is understandable with all properties of functional groups only for a person familiar with the language and paradigms of organic chemistry. It is such an important issue that the visualization techniques in the teaching of chemistry have been studied separately (Nodzyńska, 2012). At the same time, the use of visualization for the transfer of knowledge carries the risk of domination of the form over the content, or reversely, of overloading with content, which inevitably leads to the illegibility of the message. For example, a graph may carry a significant amount of information, but when the number of nodes (vertices) of the graph exceeds a certain threshold value, it easily becomes illegible, e.g. due to limited printing resolution or inefficiency of the human eye (Tamassia, 2014). This problem is experienced, for example, by analysts using the technique of classification trees or dealing with social networks. Sometimes, however, finding the optimum but difficult is possible, and advanced visualization techniques allow to merge many cognitive layers into one hybrid graphic (Opiła, 2015, Opiła, 2016a).

Case study: Visual Sentiment Analysis

Stochastic econometric models enables the analysis of the functioning of an economic entity, and testing of various control data combinations. Numerical experiments that can be performed on the model cannot be carried out in the real world due to the risk of irreversible financial damages. For example What-If sentiment analysis may be easily carried out using Characteristic Surface Model (Opiła, 2005, Opiła, 2006, Opiła et al., 2008).

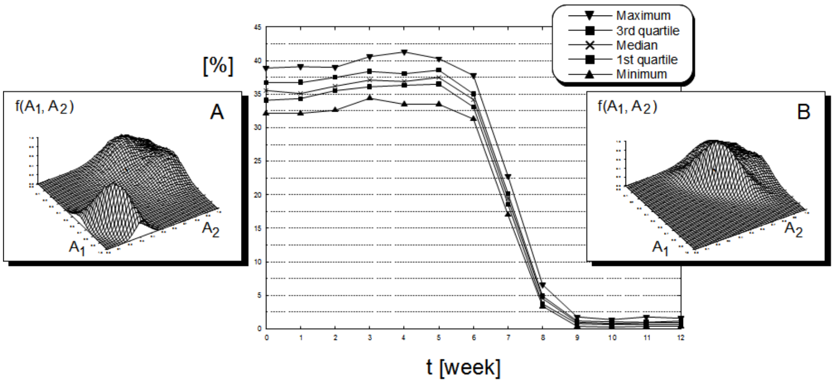
Figure 3
Evolution of the Characteristic Surface due to Change of One of Components



Source: Author's illustration, (Opiła, 2008), reworked

The essence of the model is to pay special attention to the mechanism of choosing one of the possible variants. The key element of the model is the criterion function ***K***, which determines the choice of a particular option based on the personal preferences. For simplicity, it was assumed that each individual is about to choose one of two possible options only. Assuming that a given person can intuitively assign a numerical value of attractiveness (e.g. from interval [0,1]) to both available options $A_1(O_1)$ and $A_2(O_2)$, we can mark her on a unit square by assigning coordinates (A_1, A_2). By repeating this procedure for a certain population, one obtains an empirical probability distribution. Because human preferences are characterized by considerable inertia, this distribution is slowly evolving rather than undergoing rapid changes, which is why it can be treated as an element characterizing a given population at mid-terms.

Figure 4
Estimated Financial Outcome Obtained by Means of Characteristic Surface Model



Source: Author's illustration, (Opiła, 2008), reworked

The entire population can be divided into subgroups, similarly reacting to changes in the environment. For example, for windsurfers windy, even cloudy weather is very demanded while regular tourists prefer a lack of wind and a cloudless sky. Therefore, although both will choose the date of their leave in the period of greater probability of the desired weather conditions, these will be quite different periods. The behaviour of each of the subgroups will be in response to the external stimulus different, but based on historical analysis based on Big Data, easy to model. Assuming that one of the options is to rent an apartment and the second cancellation of the trip, you can model the demand for tourist services depending on the expected conditions (weather, exchange rates, fuel prices). In turn, using the estimated demand as a coefficient in an econometric model for a tourist facility, its financial outcome for the given period may be modelled.

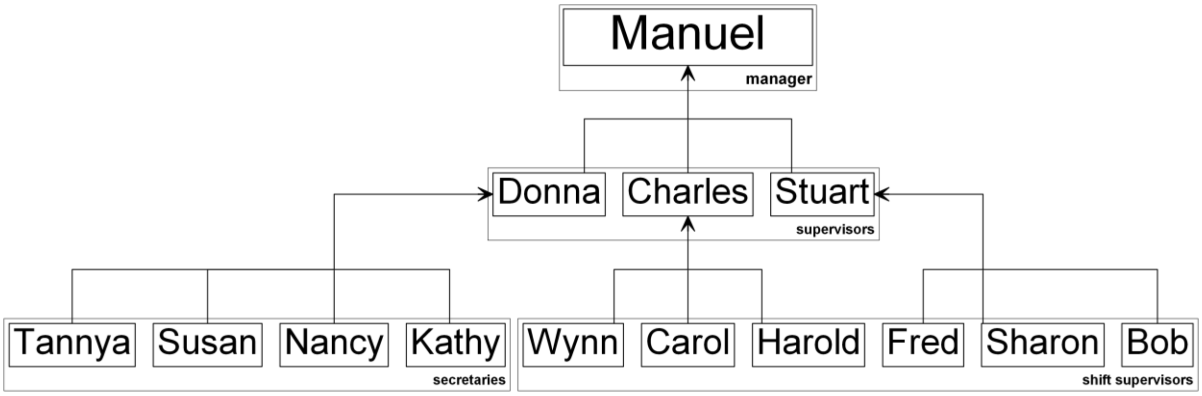
Such an analysis was carried out for a hypothetical agritourism farm (Opiła et al., 2008). The population consisted of three subgroups, one of which evolved, while the two remained unchanged. The evolution of only one subgroup had a visible effect not only on the appearance of the characteristic surface but also resulted in a drastic breakdown of the financial result (Fig. 3). The process flow can be visualized on one hybrid chart, shown in Fig. 4. The financial result obtained through Monte Carlo simulation is presented in the middle graph.

In the presented case, the process can be analyzed not only within the numerical econometric model but also on the basis of visual analysis of the characteristic surface, which makes it more understandable and useful for the average educated user.

Case study: Insight into Hidden Organizational Structure

For every organization, including business entities, steady discovery, analysis, verification and improvement of the organizational structure is an important element of the survival strategy on the increasingly competitive market. This implies searching for better and better tools to achieve this goal. In addition to proven surveys or participant observation methods, it is worth looking for methods that are independent of the level of cooperation of the surveyed employees. One of the possible approaches is an attempt to discover the actual, operational structure of the company by examining traffic in its communication networks (Opiła, 2017a).

Figure 5
Formal Organizational Structure of Investigated Hypothetical Firm

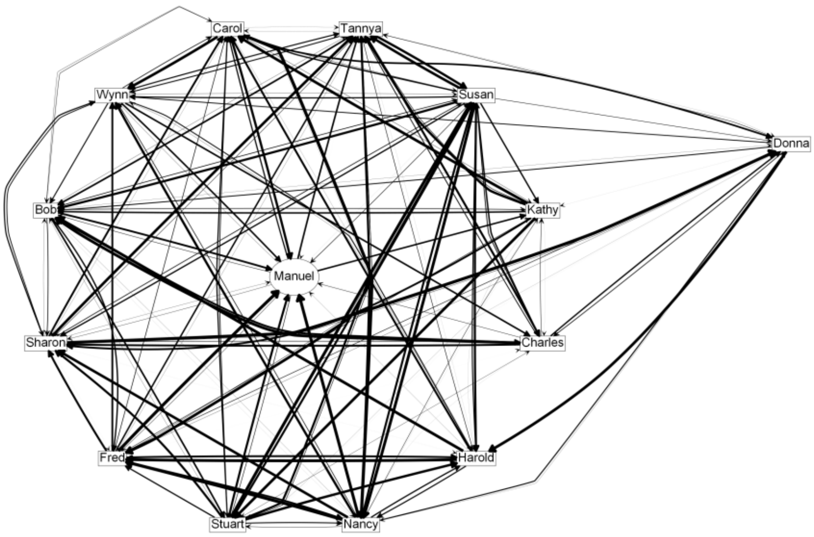


Source: (Opiła, 2017a)

The organizational structure may be defined as: „*The typically hierarchical arrangement of lines of authority, communications, rights and duties of an organization. Organizational structure determines how the roles, power and responsibilities are assigned, controlled, and coordinated, and how information flows between the different levels of management.*" (Business Dictionary, 2018). This definition emphasizes the role of communication and information transfer paths between individual elements of the organization, in particular between the management levels. The organizational structure of an enterprise, created at foundation time may undergo, more or less, evolutionary change. This may result from the intended reorganization, or from spontaneous tendency of human teams to self-organize. Thus, concealed or even secret elements of the organizational structure may arise. Their existence may solve some ailments of the official organizational structure but sometimes may pose real threat. This is why analysis of the actual organizational structure and comparing it with the intended theoretical construction is an essential element of supervision over the efficiency, effectiveness and security of the organization.

Thus, the problem arises how to discover the actual organizational structure. One can just survey employees about affiliation with a given department and the strength of this relationship according to a specially designed scale (Brzeziński et al., 2015, p.100). There, the weaknesses of the company's hitherto formal organizational structure were identified and the concept of a new, favourable to its growth was proposed (ibidem, p. 104). In this case, analysis was carried out by visual analysis based on network visualization (PAJEK program, 2018). However, the survey method has several disadvantages. First of all, it assumes an objective correctness of the answer, secondly, such a study may arouse the feeling of uncertainty in the company, and thirdly it absorbs the employee's time. It seems that analysis of communication lines and information flow paths between the nodes of the organization allows to overcome these drawbacks. In order to check this assumption, a visual analysis of a hypothetical company (Tamassia R., 2014, p.808) with the structure shown in Fig. 5, was carried out using the graphviz program (graphviz, 2018). The analysis included communication between employees via telephone or e-mail (data obtained by computer simulation in this case).

Figure 6
Number of Conversations Between 4th and 7th Hour of the Business Day Presented by *twopi* Algorithm



Source: (Opita, 2017a)

From the database one can obtain extracts containing the searched data, e.g. a number of calls during a selected time interval, say between the fourth and seventh hour of business day. The resulting data structure can be visualized using a directed graph, where the communication partakers are represented as the graph's vertices and the edge thickness of the graph reflects the examined feature, here the number of conversations. The exemplary resulting graph, constructed using the twopi algorithm, (graphviz, 2018), is shown in Fig. 6. It is clearly visible that while Manuel is the manager, Donna seems to create a separate decision centre of the company, but Charles and Stuart, who are on identical position do not distinguish from other workers. Moreover, it can be spotted, that Stuart had intense communication with Susan (secretary) while Charles communicated intensely with Bob, shift supervisor. Thus it may be supposed, that Stuart and Susan are in personal relationship but Bob consulted with Charles any production disturbances. These, and other spotted fluctuations may and should be thoroughly investigated in order to improve efficiency of the management.

Discussion

In both cases discussed, visualization plays a significant but slightly different role. The first one illustrates the numerical analysis and its conceptual basis. In the second case, it is essentially an independent tool of inference, while maintaining an illustrating role. Thus, visualization plays a supporting as well as a dominant role in the transfer of the knowledge. In each case, proper balance of visual effects (improves attractiveness) and clarity (enables understanding) is required.

Conclusion

Well-designed visualization may greatly improve efficiency and accuracy of the knowledge transfer. However badly constructed charts may negatively affect understanding of presented concepts. New phenomena may require innovative forms of graphical expression.

References

1. Brzeziński, Ł., Wyrwicka, M. (2015), „Wykorzystanie wizualizacji sieciowej do stymulacji wzrostu przedsiębiorstwa" (The use of network visualization to stimulate enterprise growth), in Kiełtyka, L., Niedbał, R. (Ed.), Wybrane zastosowania technologii informacyjnych wspomagających zarządzanie w organizacjach (Selected applications of information technologies supporting management in organizations), Publ. P.Cz., Częstochowa, p. 100.
2. Business Dictionary (2018), WebFinance, Inc., available at: https://goo.gl/hWHDdw (20 March 2018)
3. Diffen (2018), Data vs. Information, available at: https://goo.gl/9hxxgx (24 March 2018)
4. Gosztyła, K., Pulka, M., Milewicz, T., Lurzyńska, M., Opiła, J., Stochmal, E., Hubalewska-Dydejczyk, A., Galicka-Latała, D., Krzysiek, J. (2011), „The impact of polychlorinated biphenyls on placental and ovarian steroidogenesis", Przegląd Lekarski, Vol. 68 No. 2, pp. 118–122.
5. graphviz (2018), computer program, available at: http://www.graphviz.org/ (22 March 2018)
6. Kozlíková B., Krone M., Falk M., Lindow N., Baaden M., Baum D., Viola I., Parulek J., Hege H.-C., (2017), "Visualization of Biomolecular Structures: State of the Art Revisited", Computer Graphics Forum, Vol. 36 No. 8, pp. 178–204.
7. Kuźmińska-Surowaniec, A., Pawelczak, I., Michałowicz, J. (2016), „Eliminacja bisfenolu A z wody i ścieków z wykorzystaniem metod fizyko-chemicznych i enzymatycznych" (Elimination of bisphenol A from water and sewage using physico-chemical and enzymatic methods), in Garbacz, J.K. (Ed), Diagnozowanie stanu środowiska: metody badawcze – prognozy, kompleksowe badania i ochrona środowiska naturalnego (Diagnosing the state of the environment: research methods -- forecasts, comprehensive research and protection of the natural environment) BTN, Bydgoszcz, pp. 73-80.
8. Nodzyńska, M. (2012), Wizualizacja w chemii i nauczaniu chemii (Visualization in chemistry and teaching chemistry), UP Kraków, Kraków.
9. Opiła, J. (2005), „Analiza ewolucji nastrojów konsumenckich z wykorzystaniem metody powierzchni charakterystycznej" (An analysis of the evolution of consumer sentiment using the characteristic surface method), in Duda, J.T. (Ed.), Systemy informatyczne i metody obliczeniowe w zarządzaniu (IT systems and calculation methods in management), AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków, pp. 141–148.
10. Opiła, J. (2006) „Analiza wrażliwości stochastycznego modelu ekonometrycznego z zastosowaniem rozszerzonej metody powierzchni charakterystycznej" (Analysis of the sensitivity of a stochastic econometric model with the use of the extended characteristic surface method), in Duda J.T., Waszkielewicz W. (Ed.), Nowoczesne metody i techniki w zarządzaniu (Modern methods and techniques in management), AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków, pp. 200-210.
11. Opiła, J., Opiła, C., Cichowska, J. (2008) „Prognozowanie ostrzegawcze w mikroprzedsiębiorczości z zastosowaniem rozszerzonej metody powierzchni charakterystycznej na przykładzie gospodarstwa agroturystycznego" (Emergency

forecasting in microenterprise with the use of the extended surface method on the example of an agritourism farm), in „Zarządzanie przedsiębiorstwem – teoria i praktyka" (Business management - theory and practice), XI Międzynarodowa Konferencja Naukowa, Kraków, WZ AGH, pp. 1-10.
12. Opiła, J. (2015), "Prototyping of visualization styles of 3D scalar fields using POV-Ray rendering engine", MiPro 2015: 38th International Convention on Information and Communication Technology, electronics and microelectronics, May 25-29, 2015, Opatija, Croatia, MIPRO Croatian Society, Opatija, Croatia, pp. 328–333.
13. Opila, J. (2016a), "Prototyping of visualization designs of 3D vector fields using POVRay rendering engine", MiPro 2016: 39th International Convention on Information and Communication Technology, electronics and microelectronics, May 30 – June 3, 2016, Opatija, Croatia, pp. 343-348.
14. Opiła, J. (2016b), „Wizualizacja szlaków migracyjnych zwierząt z użyciem biblioteki ScPovPlot3D" (Visualization of animal migration routes using the ScPovPlot3D library), in Garbacz, J.K. (Ed.), Diagnozowanie stanu środowiska: metody badawcze – prognozy, kompleksowe badania i ochrona środowiska naturalnego (Diagnosing the state of the environment: research methods -- forecasts, comprehensive research and protection of the natural environment), BTN, Bydgoszcz, pp. 81-92.
15. Opiła, J. (2017a), „Wizualna analiza struktury organizacyjnej z wykorzystaniem narzędzi Open Source" (Visual analysis of the organizational structure with the use of Open Source tools), in Kiełtyka, L., Sokołowski, A. (Ed.), Techniki i technologie wspomagające funkcjonowanie przedsiębiorstw (Techniques and technologies supporting the functioning of enterprises), Publ. P.Cz., Częstochowa, pp. 191–203.
16. Opiła, J. (2017b), „Wizualizacja jako narzędzie transferu wiedzy" (Visualization as a Knowledge Transfer Tool), in Garbacz, J. (Ed.), Diagnozowanie stanu środowiska: metody badawcze - prognozy: kompleksowe badania i ochrona środowiska naturalnego: zbiór rozpraw (Diagnosing the state of the environment: research methods - forecasts, comprehensive research and protection of the natural environment), BTN, Bydgoszcz 2017, pp. 169-182.
17. PAJEK (2018), computer program, available at: http://mrvar.fdv.uni-lj.si/pajek/ (22 March 2018)
18. Tamassia, R. (2014), Handbook of Graph Drawing and Visualization, CRC Press, Boca Raton FL

About the author

Janusz Opiła, Ph.D. is an Adjunct Professor at the Faculty of Management, AGH-UST, Kraków, Poland, in Department of Applied Computer Science (ORCID 0000-0003-1179-1920). He received Ph.D. in Physics at the Faculty of Physics and Applied Informatics of AGH-UST with the dissertation thesis concerning usage of Boundary Element Method for computation of magnetic fields. His main research interests are information visualization and photorealistic visualization. He is actively engaged in a legal government project. He published several scientific papers in international and national journals, co-authored two books and participated in several scientific international conferences. The author can be contacted at **jmo@agh.edu.pl**.