

Semiotic Assessment of Visualization Tools

Alena Zakharova¹, Aleksey Shklyar² and Evgeniya Vekhter²

¹ Institute of Control Sciences of Russian Academy of Sciences, 65, Profsoyuznaya st., Moscow, 117997, Russian Federation

² Tomsk Polytechnic University, 30, Lenin Ave., Tomsk, 634050, Russian Federation

Abstract

A visual analytics means is a set of tools used to obtain a visual representation of the data explored by the user, its interpretation and complete analysis. It is quite possible that such a set will be a developed specialized environment for continuous user interaction with a data stream, which has the capacity to control and replace the way of visual data representation. A wide range of existing visualization tools, as well as unlimited borrowing of technological solutions originally intended for a narrow application area, and the transfer of the capabilities they create to other directions make it urgent to form a sound approach to the comparison and selection of means of data visual representation of a given complexity [1]. The purpose of this selection is to save the end user's resources, based on the absence or minimization of resource losses arising at the stage of selection and testing of visual analytics tools. The paper shows the possibility of using a semiotic visualization model for purposeful design or selection of visualization tools.

Keywords

Visualization, visual analytics, semiotic model, interpretation

1. Selection problem statement

The visualization tool and visual analytics selection problem arises only in matters of their use by the user. However, the requirements usually imposed on tools for solving visual analytics tasks are as follows:

- accuracy and completeness of the analysis performed;
- resource intensity and efficiency of the solution process;
- the possibility to obtain additional useful results of user interaction with data.

Issues related to understanding the correspondence of the available visualization tools to the purpose of their application in a specific task, as well as the differences in the use of these tools by different people, remain extremely complicated. If each available visualization tool leads to an analysis result that meets specified requirements, then the selection of visual analytics tools can be determined by subjective advantages: convenience, emotional effect, compliance with individual user restrictions (speed of work, familiarity or comfort of the working environment, the need for additional training etc.). Difficulties arising at this stage of research or another practical activity can have a significant impact on the final result, and the amount of resources spent does not always provide the user with an appropriate level of result [2]. To clarify this statement, let us consider a few common situations.

Example 1. Selecting a visualizer in a designer's work. When a practicing designer determines the visualization tools that ensure the commercial success of the project being executed, it is necessary to determine the boundary conditions of the visualization task, one of which is the quality level of the final images. The most important problem for every designer is that there is no way to accurately indicate the desired features of the visualizing result. Photorealism, often proposed as a quality criterion, is highly controversial in design operating fictional objects.

GraphiCon 2021: 31st International Conference on Computer Graphics and Vision, September 27-30, 2021, Nizhny Novgorod, Russia

EMAIL: zaawmail@gmail.com (A. Zakharova); shklyarav@tpu.ru (A. Shklyar); vekhter@tpu.ru (E. Vekhter);

ORCID: 0000-0003-4221-7710 (A. Zakharova); 0000-0003-4442-7420 (A. Shklyar); 0000-0003-0604-0399 (E. Vekhter)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

Moreover, common examples represent situations in which the hyperrealism of visualization of spaces saturated with objects, requiring high image detail, is the cause of information overload, false perception, and an uncontrolled shift in the user's focus of attention. In addition, a high degree of accuracy in the details of the image and the lack of balance between significant and minor elements become sources of cognitive delusions, when the search and assessment of the relevant dependencies in the visual image by the observer leads to the formation of erroneous interpretations. Their verification requires additional resources and can be considered both a useful and undesirable process.

The negative result can include inadvertently misleading the viewer, creating inappropriate associations, or creating unreasonable expectations. In the example under consideration, the visualization problem is associated with the impossibility of distinguishing between true and false ideas about the object of perception, since it is just a difference in the perception of the viewer and the author.

Example 2. Selecting scientific visualization tools. Visual data representation can be a way to compare the results of computational experiments describing the same process but obtained, for example, using different models. A problem is caused by the situation when the visualized calculation results cannot be evaluated due to the lack of a reference example or the user's understanding of the validity criteria. Thus, visualization can be useful in the following cases: comparison of broad data pictures if the data structure is difficult to quickly evaluate; search for differences in data related to different sources if the differences are small or unpredictable (cannot be estimated algorithmically).

Thus, when choosing visual analytics tools, it is necessary to compare not the accuracy and speed of translating the initial data into a visual representation, which often act as evaluation criteria, but, for example, the possibility of using the visual representation metaphor, which allows (subject to other boundary requirements) achieving the purpose of data analysis most efficiently. It is commonly known that this purpose is the user's understanding of the meaning of these data.

2. Semiotic comparison

If the metaphor necessary for the efficient visual data interpretation is known (developed, verified, the method of application is prepared), then the visual analytics tool selection is determined by the advantages of the metaphor itself. In the semiotic visualization model [3], this statement corresponds to the choice of the optimal language of visual representation, serving the purpose of visualization. In a broader context, the semiotic visualization model which establishes correspondence between visual perception and information communication allows for effective analogies with basic linguistic definitions. As a result, there is a useful opportunity to develop visual analytics tools that differ in their parameters depending on the content and characteristics of the communication process. Consequently, the advantage of visualization tools can be an easy-to-change, wide-ranging possibility of achieving an exact match to the user's actual goals. Figure 1 shows a variant of the semiotic model demonstrating the need for a mandatory and precise agreement between the capabilities of visualization tools and their application. The lack of such an agreement, despite the high-quality level of visualization, will lead to the solution of a problem that is different from the one posed. The question is how to achieve this conformance in the most efficient way? By what means? At what stage of visual research?

It is necessary to answer the question of what makes one image more understandable for the viewer than another one? As part of the discussion of the capabilities of visual analytics, at the initial stage, this question can be formulated differently: what properties of the image help the viewer to complete the interpretation, i.e. to formulate, test and accept the hypothesis of interpretation more efficiently (in most situations, faster) than in their absence? Reliance on a semiotic visualization model allows answering this question by pointing out the dependence of the effectiveness of a particular visualization tool on the interpretation purpose:

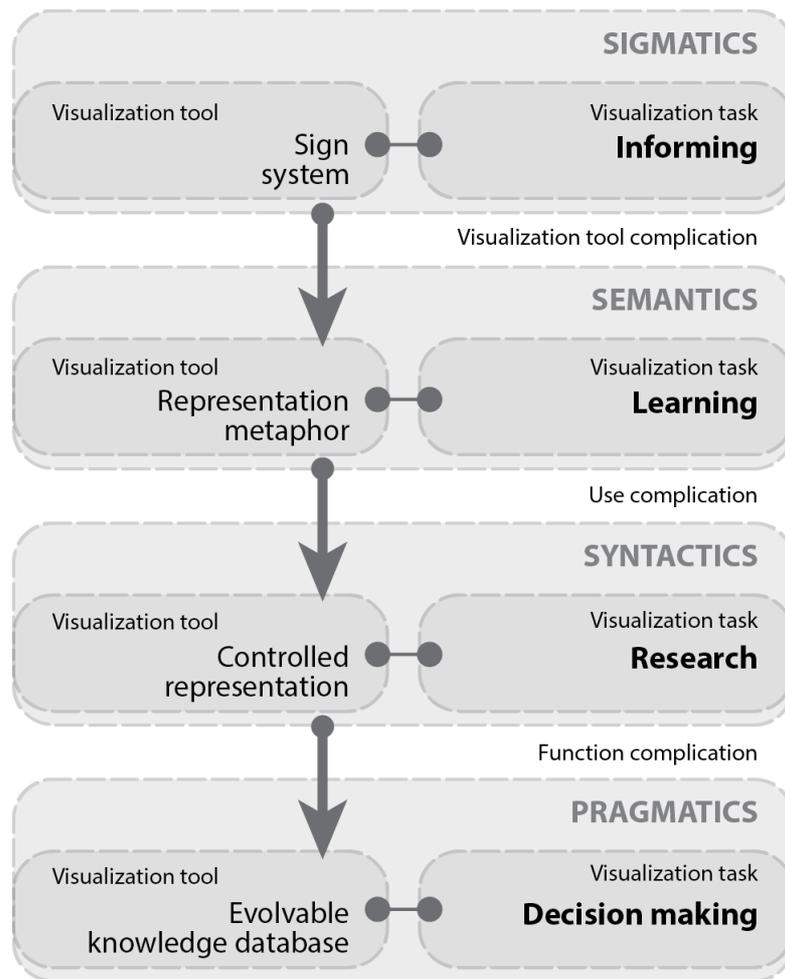


Figure 1: Implementation of the semiotic model in the development of visual analytics tools

- The task of informing. For situations in which a visual data image is intended to convey to the viewer information that does not require a discussion, the image must contain a visual statement formulated in a known visual language (expressive means familiar to the user) and disallow interpretation variants.
- The task of learning. In this case, the aim is to change or supplement the user's own knowledge. Therefore, visual analytics efficiency in solving the learning problem presumably depends on the conflict of the user's "old" knowledge system and visual information involved in the learning task [4]. Capabilities of visualization tools become user-dependent. In this case, the increase in the visual analytics tool efficiency depends on the immediate determination of user knowledge that prevents new experience and the availability of adaptation capabilities of visual analytics tools.
- The task of analysis (research). The visual picture should (one of the goals) create freedom for the appearance of options for interpretation hypotheses. Perhaps, by analogy with the linguistic model, a visual statement may possess "understatement", i.e. offer (direct) the viewer to supplement the image with his own meanings. The main task of visual analytics tools is to help the researcher in obtaining variants of hypotheses and in choosing the one that is closer to the individual cognitive worldview. Thus, the possibility of visual search is the most interesting when comparing visual analytics tools.
- The task of decision making. It is the most difficult task. Here, the user of visualization tools makes a choice in favor of one of the possible scenarios, the starting point of which is the result of the visual image interpretation [5]. One of the goals of visualization is the user's independent selection of decision-making rules. In many applied cases, selection difficulties are described by at least two reasons:

- 1) lack of a methodology for comparing options or a comparison scale. It is typical for multidimensional or heterogeneous data [6];
- 2) user's doubts resulting from the similarity of the compared options and incomplete understanding of the role of differences after the final decision.

In applied problems, various ways of visual representation of any data can manifest themselves in different ways. Therefore, when choosing visual analytics tools, first of all, knowing the parameters of the application area, it is necessary to determine the requirements for visualization tools. In accordance with the considered semiotic visualization model, the possibility of using visual analytics tools in tasks other than the area of their initial application appears after introducing the necessary restrictions or additions. For example, when the freedom of interpretation is limited, visual learning means move into the category of informing means and acquire the corresponding applications. The converse statement is also true, which consists in the fact that purposeful visualization tool selection changing the way of communication with the user creates for him an externally controlled interpretation of the data received. Thus, the semiotic visualization model becomes the basis for the development of methods for manipulating the information communication participants' perception and understanding.

3. Effectiveness or persuasiveness

The study of the semiotic visualization model makes it possible to indicate not only the continuity of some characteristics of visualization tasks and their sequential complication, but also a certain deep contradiction resulting from the peculiarities of visual perception. With this in mind, it is necessary to highlight two potential advantages of visualization [7] which work in all cases but in different ways.

- First, visual representation is used to detect errors or gaps in data under study and to find inconsistencies between actual and expected results. A significant number of visualization techniques have been developed that allow achieving this goal quickly and efficiently due to the evolutionarily developed mechanisms of visual perception.
- Secondly, visualization transforms the initial data, as an unconscious abstraction, into a perceived form that is offered for interpretation. In most cases of its application, visualization has a high degree of reliability for the viewer, and in the conditions of formation of a digital society and a new culture of perception, there appears a technology of influencing the user which has unlimited persuasion potential. As follows from the semiotic model, in different combinations of expressive means, this corresponds to the solution of different visualization problems.

Thus, visualization technology development and high speed of perception lead to the emergence of information-rich images, which are the basis of visual analytics systems [8]. In turn, this generates the user's deformed perception. It is necessary to clarify for him the visualization task and the possibility to control the visualization tools.

To assess the influence of this contradiction on the process of interaction between the user and the visual image, a series of tests has been carried out with the participation of design students. In each test, the participants were simultaneously shown three images of the same content: a photograph and results of 3D modeling using several common visualizers (Corona Render, V-Ray). It was proposed to determine the real image (selection), indicate the inaccuracies of visualization (errors), approve the demonstration method (content assessment). The series differed in the level of information saturation of the images. A simple image represented a geometric shape without environment with two light sources. The results of the tests are shown in Table 1.

Table 1
Simple image perception

Question	Photo	Visualizer 1	Visualizer 2
Selection	68%	22%	10%
Errors	8%	18%	12%
Content assessment	36%	72%	61%

In a series with a complex image, an image of an interior with natural light was demonstrated (Figure 2). However, after several trial tests, the interior photograph was replaced with an image obtained using another renderer (Mental Ray). The reason for this was the need to reduce the role of the reference element in the overall analysis. Therefore, the selection criterion corresponded to the participant's definition of the most correct image from a subjective point of view. The results of the tests are shown in Table 2.

Table 2
Complex image perception

Question	Visualizer 1	Visualizer 2	Visualizer 3
Selection	32%	24%	44%
Errors	36%	42%	28%
Content assessment	48%	32%	20%

As a result, preliminary conclusions have been obtained that correspond to the semiotic model (Figure 3) and the previous reasoning about the contradictions in visualization:

- Increasing saturation of the image with additional elements can lead to an erroneous interpretation.
- Experience with visualization tools can improve the speed and error rate found in the rendered data.
- The use of expressive visualization tools is a convenient tool for managing attention and interpretation of the analyzed image.
- In the absence of a reference standard, users independently develop the criteria they need.
- The necessity and possibility of careful visual analysis can change the initial assessment.
- Erroneous hypotheses generated by the observer can be rejected by examining additional image elements.



Figure 2: Interior visualization options involved in a series with a complex image

4. Requirements for visual analytics tools

In the development of visual analytics tools, it is a huge challenge to formulate even the most general requirements for visualization tools to ensure their efficiency. Let us consider one of the common tasks associated with the use of visualization tools for immediate comparison of two or more data sets in order to determine their reliability and select the most preferable one [9]. In the general case, there is no comparison scale; there is no formal standard for the baseline assessment. In other words, there are only sets of compared data at the researcher's disposal, and external requirements or formulated reliability criteria are not enough to unambiguously determine an efficient approach to the use of visual analytics tools.

Approach to the task of visual informing. Independent data representation to the user for their comparison is possible only at the very first step – when the user is provided with the initial data. As the test studies have shown, in the sequential presentation of visualized data, the assessment and interpretation of the second and subsequent sets is influenced by the perception and interpretation of all

previous data. In the case when independence of judgment is important for general comparison of efficiency of visual analytics tools, it is necessary to reduce the “first encounter effect” when the first sample is studied in detail and the subsequent ones only in comparison with the previous one. Techniques that provide unambiguous informing of the user include strict perception timing (the rhythm of data flow) or any other visualization method focused on limiting interpretation options (for example, providing data in a form for memorization, not understanding, without highlighting features, internal dependencies, etc.). In a test case with comparing the work of visualizers, this can be implemented when demonstrating fragmented images, with emphasis on individual elements, with a large amount of detail but without forming a complete subjective impression by the viewer from the full visualization options. Within the semiotic model, visual information tools are used for the user to accept the choice as an obvious fact.

In the approach to the task of visual learning, the goal is to inform the user about the differences, i.e. the advantages and disadvantages of one dataset over another one. In this case, the principle of distinguishing differences makes it unnecessary for the user to obtain a full information picture. In this approach, the user is persuaded of the correctness of the conclusions which are the basis of the visual message. Here, the possibilities of expressive means involved in visualization are aimed at limiting independent conclusions and emphasizing the necessary ones. For example, a beautiful visualization of an interior becomes preferable to a physically correct one. In this case, the artificial selection formed due to the visualization shapes or complements the user's cognitive worldview, but at the same time it does not have to be correct in general understanding. Consequently, visual analytics tools that operate within the learning task provide selection accompanied largely by external arguments.

In the research task, the main goal of visual analytics tools is to find arguments that the user lacks in order to understand the reasons for the differences between the compared sets and their role. To organize the search procedure, it is necessary to have a request (waiting) and the ability to form a response (feedback). The absence of relevant elements in visual analytics tools equates the conditions to the task of informing (or learning).

A passive (arising as a part of the initial data, as a result of the formal execution of the visualization procedure) request perceived by the user as the goal of the research can be presented as a contradiction (in the user's understanding) or incompleteness in the visual image. An alternative option, or an active request, is a user's meaningful waiting seeking to find in the visual image a confirmation (refutation) of an already existing interpretation hypothesis. In this case, consistency between the language means related to the user's thinking and the visualization metaphor becomes mandatory. The required consistency is achieved in two ways: user's learning (familiarization) or selection (adjustment) of the visualization metaphor already studied earlier. The reasons for differences in the interpretation of the question can be associated with linguistic, emotional, cultural, physical factors that determine communication participants' state and capabilities), as well as the user's local (in time) awareness about the origin and characteristics of the visualized information. Thus, interactive control of visual analytics tools and the process of their use becomes the main element in solving the research problem [10].

In the task of decision making, there is a search for options, their comparison and selection of the one most suiting a goal or a set of related goals of communication with data. Therefore, for such a purpose, replacing the visual representation of the initial data with the visualization of options generated by formal rules can be a way to optimize the solution [11]. If formal rules for obtaining selection are not defined, then the process of visual communication is divided into research and evaluative categories. In some practical situations, the problem of decision making can be reformulated into the problem of assessing the consequences of the decision and selection based on the predicted results [12]. For such a formulation of the research goal, visualization can be convenient due to the clarity and persuasiveness of the visual representation of data that already appear in the process of visual research, i.e. are not initial data.

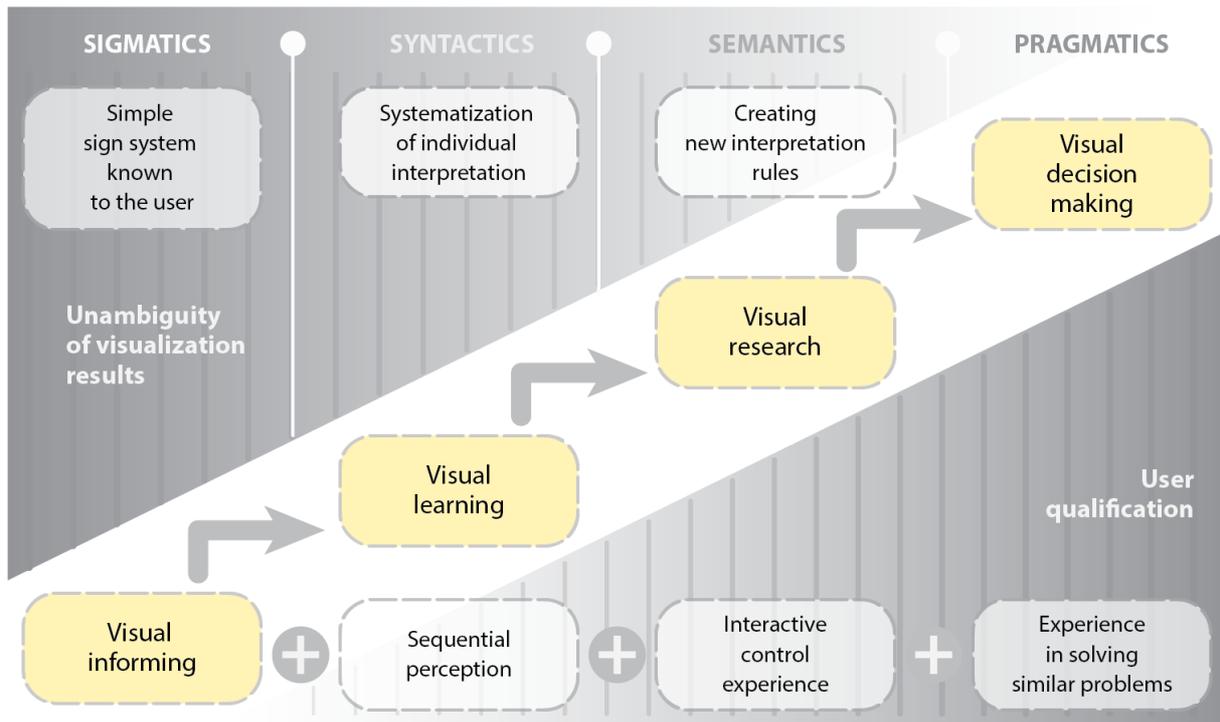


Figure 3: Correlation between the unambiguity of information visual representation and the user's preparedness for various visual analytics tools

5. Conclusion

The initial visualization task changes in many application situations. On the one hand, the initial data which are the source of visualization are replaced or supplemented with new ones which are the result of intermediate stages of the user's communication with the data. On the other hand, there is selection and subsequent refinement of the visualization method, the features of which are determined by the user's capabilities. It should be noted that the correction of the visualization method can be significant, changing in the process of visual communication, and can also act as a technique that stimulates the user's cognitive activity.

The paper considers the approach to assessing or formulating the requirements for the designed visual analytics tools based on the semiotic model which makes it possible to change the procedures for visualizing and designing the corresponding tools. Supplementing visual analytics tools with the possibility of flexible and directed interaction with the user becomes a resource for increasing their efficiency due to the possibility of timely transition between different types of visual research tasks. Thus, the paper proposes an approach to comparing visual analytics tools based on the need to jointly analyze the purpose of data exploring, the features of visualization tools and the user's individual characteristics. A change in each of these components, considered due to the proposed semiotic visualization scheme, is a solution if it is necessary to use visualization tools with increased effectiveness.

6. References

- [1] A.A. Zakharova, E.V. Vekhter, A.V. Shklyar, The Applicability of Visualization Tools in the Meta-design of an Educational Environment. *European Journal of Contemporary Education* 8 (1) 43–51 (2019). doi: 10.13187/ejced.2019.1.43.
- [2] M. Chen, A. Golan, What May Visualization Processes Optimize? *IEEE Transactions on Visualization and Computer Graphics* 22(12) 2619–2632 (2019). doi: 10.1109/TVCG.2015.2513410

- [3] A.A. Zakharova, E.V. Vekhter, A.V. Shklyar, Adaptable Visualization. *Scientific Visualization* 13 (4), 67–78 (2021). doi: 10.26583/sv.13.2.05
- [4] C. Vieira, P. Parsons, V. Byrd, Visual Learning Analytics of Educational Data: A Systematic Literature Review and Research Agenda. *Computers & Education* 122, 119–135 (2018). doi: 10.1016/j.compedu.2018.03.018
- [5] L.V. Massel, A.G. Massel, R.A. Ivanov, Integration of Visual Analytics, Cognitive Graphics and Semantic Modeling in Semiotic Intelligent Systems for Decision Support, in: Proceedings of International Workshop "Contingency Management, Intelligent, Agent-based Computing and Cyber Security in Critical Infrastructures", Irkutsk, MESI SB RAS, 2016, pp. 19-22.
- [6] H. Chen, et al., Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration. *IEEE Transactions on Visualization and Computer Graphics* 21(3), 1072–1086 (2015). doi: 10.1109/TVCG.2015.2410278
- [7] A. Dasgupta, et al., Familiarity Vs Trust: A Comparative Study of Domain Scientists' Trust in Visual Analytics and Conventional Analysis Methods. *IEEE Transactions on Visualization and Computer Graphics* 23(1), 271–280 (2017). doi: 10.1109/TVCG.2016.2598544
- [8] G.K.L. Tam, V. Kothari, M. Chen, An Analysis of Machine- and Human-Analytics in Classification. *IEEE Transactions on Visualization and Computer Graphics* 23(1), 71–80 (2017). doi: 10.1109/TVCG.2016.2598829
- [9] A.E. Bondarev, On the Construction of the Generalized Numerical Experiment in Fluid Dynamics, *Mathematica Montisnigri XLII*, 52–64 (2018).
- [10] A. Batch, N. Elmqvist, The Interactive Visualization Gap in Initial Exploratory Data Analysis. *IEEE Transactions on Visualization and Computer Graphics* 24(1) 278–287 (2018). doi: 10.1109/TVCG.2017.2743990
- [11] A.G. Podvesovskii, R.A. Isaev, Constructing Optimal Visualization Metaphor of Fuzzy Cognitive Maps on the Basis of Formalized Cognitive Clarity Criteria. *Scientific Visualization* 11 (4), 115–129 (2019). doi: 10.26583/sv.11.4.10
- [12] V.I. Averchenkov, V.V. Miroshnikov, A.G. Podvesovskii, D.A. Korostelyov, Fuzzy and Hierarchical Models for Decision Support in Software Systems Implementations, in: A. Kravets et. al. (Eds.), *JCKBSE 2014*, volume 466 of Communications in Computer and Information Science, Springer Cham, 2014, pp. 410-421. doi: 10.1007/978-3-319-11854-3_35