Integration of interdisciplinary and evidence-based approach into research policy

To cite this article: B I Savelyev *et al* 2022 *IOP Conf. Ser.: Earth Environ. Sci.* **1069** 012041

Integration of interdisciplinary and evidence-based approach into research policy

B I Savelyev[1], D V Polevoi[1] and S V Pronichkin[2]

[1]Federal Center of Theoretical and Applied Sociology of Russian Academy of Sciences, 24/35, Krzhizhanovsky Street, Moscow, 117218, Russia
[2]Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 40, Vavilov Street, Moscow, 119333, Russia

E-mail: savelyevbi@mail.ru

Abstract. Sustainable development scientific concepts can be presented in the form of an integrated hierarchical network of frames. A knowledge base ontology model has been developed, which allows calculating the importance of frames, due to which the mechanism of adaptation to a given interdisciplinary field of scientific research is implemented. The ontology is identically represented as a multiplexed semantic network. An algorithm for determining the relevant frames based on the multisets metric spaces multiplicities has been developed. Mathematical and software support for the functioning of the decision support system has been developed, depending on the class of tasks for ensuring sustainable development.

1. Introduction

The globalization aspects of the modern society development necessitate the construction of complex systems for the functioning of individual subject areas. Using the example of sustainable development of complex socio-economic systems, this is the determination of critical indicators of the territories development on the basis of semi-structured big data. However, this is difficult to do due to the discrepancy between the requirements for information systems and the need to organize (search for objects, their systematization, coordination, data integration) of various types of information objects into a complex information system, which manifests itself through: weak structuring of links between objects; the need to include new objects in the system, the impossibility of systematization due to the large number of objects and their different nature [1, 2].

The development of information resources using various data models, control schemes requires the development of a unified method of access to them in order to enable the user to choose an adequate toolkit for studying and using various approaches to processing big data [3, 4]. The need for this arises in organizations whose work is to develop a large number of different types of data sources, for which not all semantic relationships are known and indicated. In some cases, semantic links are unknown due to an undefined number of initial sources or due to a lack of qualified people in defining such links. In other cases, not all semantic links are required to classify services to users. Therefore, users do not have a single scheme according to which they can create requests for targets.

The intelligent decision support systems technology is one of the most developed areas of artificial intelligence. Research in this area consists in the development of automated information systems that are used in those areas of human activity that require the integration of an interdisciplinary and evidence-based approach to research policy to ensure sustainable development [5, 6].

2. Ontological modeling of decision support in intelligent systems

The decision support ontology in intelligent systems can be divided into two subclasses. There is a class of tasks for which sustainable development indicators are essential [7, 8]. Such tasks are indicative. For another class of problems, the meaning of concepts is not important, but rather their semantics or the frequency of terms in the text is important. These include the clustering of scientific and technical resources, the scientific research classification, intelligent search engines, abstracting and annotation of scientific publications in the field of sustainable development. Such tasks are semantic.

For the effective functioning of intelligent decision support systems, it is necessary to build criteria on the basis of which the relevance of states or frames should be determined. The construction of such criteria directly depends on the class of tasks (semantic or indicative).

Different directions of decision support in intelligent systems require two different functional models (search for relevant frames and planning of activities), and depending on the type of use cases, the use of two different metrics for solving decision support problems and assessing the quality of the decisions obtained.

The knowledge base ontology adaptation efficiency to the subject area features is determined by the elements and mechanisms of adaptation inherent in its structure through self-learning during operation. One of the approaches to the implementation of such mechanisms is the automatic weighing of knowledge base concepts and semantic relationships between them during self-learning. This role is assumed by the coefficients of the importance of concepts and connections [9].

The importance coefficient of a concept (connection) is a numerical measure that characterizes the importance of a certain concept (connection) in a specific subject area and dynamically changes according to certain rules during the system operation. It is proposed to expand the ontology concept by introducing into its formal description the coefficients of significance of concepts and relations. Therefore, it is proposed to define the decision support ontology in intelligent systems using: concepts; relationship between concepts; interpretation of concepts and relationships that establish the semantic limitations of the system of concepts and relationships; the importance of concepts; the importance of relationships.

The ontology thus defined is adaptive, since it adapts to the modification of concepts, the coefficients of the importance of these concepts and the connections between them. Such an ontology is unambiguously represented as a multiplexed semantic network. Therefore, it is proposed to construct the metric using the weighted multiplicity indices of multisets according to decision frames [10].

Framed inference is effective when the main source of knowledge is experience, not theory. Solutions are not unique to a specific sustainability challenge and can be used elsewhere. The goal of solving a problem is not to get a guaranteed correct solution, but the best possible one.

Inference based on frames is a method of constructing intelligent decision support systems that make decisions on a given problem or situation based on the results of searching for analogies stored in the frame base [11, 12]. Such a use case is relevant. From a mathematical point of view, among the elements of the set of precedents $F = \{f_1, f_2, \ldots, f_M\}$ the relevant f^* is the frame for which the distance to the current situation A is the minimum:

$$f^* = \arg\min_{i,j} D_{ij}(f_i, A). \quad (1)$$

The process of planning activities should achieve the socio-economic system sustainability target state [13, 14]. First of all, it is necessary to build a plan for achieving this state with all possible alternatives. The planning process is based on decomposition. The planning problem contains three components: a set of states, a set of actions and a set of target states.

Consequently, to support the adoption of managerial decisions, it is necessary to evaluate states and actions on the basis of a system of criteria. The criteria are used to assess the relevance of frames as well as to assess the relevance of states. The efficiency of the functioning of the decision support system directly depends on the method of determining the criteria.

3. Criteria for evaluating the frames relevance

The construction of criteria for assessing the relevance of frames in the tasks of finding the target state should be based on ontologies. The effectiveness of assessing the relevance of interdisciplinary research frames and states of sustainable development is due to the following factors:

- transition from a non-deterministic mode of production and use of knowledge by subjects of the scientific process to a mode of effective knowledge management (at all stages of their life cycle) and science as a whole;
- the fundamental possibility of solving interdisciplinary problems of high complexity to ensure sustainable development;
- unproductive labor costs of scientists will be significantly reduced due to the direct use of sustainable integrated scientific knowledge in a canonical form;
- increase in the effectiveness and quality of solving the ensuring sustainable development problems based on an objective assessment and control of the scientific research results;
- sustainable knowledge will become intellectual capital, and the science subjects will become direct participants in the economic activity of society, which will create the most favorable conditions for stimulating the development of both science itself and a creative society.

It is proposed to determine the distance between frames and the situation as the sum of the distances of the multiset multiplicity between the most important concepts of the frame and the current situation. The most important is the concept, which is the center of various criteria. Such a center is a concept, the average distance from which to all other concepts is minimal. Obviously, the distance thus determined will depend on how the distance between two adjacent frames is determined. For this, it is proposed to use the distances between frames based on four criteria for the Hamming metric m.

$$D_{i1}(f_i, A) = \left(m(f_i \Delta A)\right)^{1/p}, \tag{2}$$

$$D_{i2}(f_i, A) = \left(\frac{m(f_i \Delta A)}{m(\overline{f})}\right)^{1/p}, \tag{3}$$

$$D_{i3}(f_i, A) = \left(\frac{m(f_i \Delta A)}{m(f_i \cup A)}\right)^{1/p}, \tag{4}$$

$$D_{i4}(f_i, A) = \left(\frac{m(f_i \Delta A)}{m(f_i + A)}\right)^{1/p}. \tag{5}$$

Further, using the multiplex semantic network that defines the frame ontology, the distance from the given frame f_i to the current situation A is determined. If the concepts of the current situation are not included in the multiplex semantic network, then the ontology of this frame is supplemented by the ontology of the entire interdisciplinary research that includes this frame. If the necessary concept is no longer included in the ontology, then its absence determines the large distance D_{ij} of the frame from the current situation.

Each interdisciplinary sustainable development project is characterized by resource costs and implementation timelines. Information about alternatives and resource costs should be stored in the ontology. Information about the meaning of the criteria is stored in the knowledge base. Obviously, new alternatives may appear, so it is necessary to constantly monitor new scientific directions and add them to the ontology [15, 16].

After evaluating actions and states, the problem of choosing a path in a multiplex semantic network is reduced to the problem of asynchronous dynamic programming [17, 18]. The solution is found in the form of a sequence of nodes for the transition from the initial to the final state.

Since the ontology forms a taxonomy of concepts, and each concept is a frame in the multiplexed semantic network. An algorithm has been developed for determining the coefficients of the importance of frames and links in the process of functioning of an intelligent decision support system for indicative and semantic tasks of ensuring sustainable development. The algorithm consists of the following main steps:

- the ontology frame importance is equal to the sum of the subframes importance and the adjacent frames importance in the multiplexed semantic network;
- after adding a new subframe, its importance and the importance of all parent frames up to the root increases by the subframe weight value;
- when establishing a connection between frames and the corresponding nodes of the ontology multiplexed semantic network, an edge appears, and the re-establishment of connections leads to the appearance of multiple edges in the network;
- multiplicity of edges reflects the frequency of communication of a pair of semantically related frames, but multiple edges after recalculation do not increase the valency of a node in the multiplex semantic network;
- the weight of a knowledge base instance is equal to the total weight of its frames in the multiplexed semantic network.

Thus, the knowledge base ontology model allows calculating the importance of frames in the process of adding, retrieving and using them in the operation of the decision support system, due to which the mechanism of adaptation to a given interdisciplinary field of scientific research in the field of sustainable development is implemented.

The coefficients of the importance of frames and the links between them are the multisets multiplicities of various interdisciplinary research topics defined in a specific area of sustainable development. If the task of classifying scientific projects is considered, then within the framework of the ontology several different interdisciplinary topics for ensuring sustainable development should be described.

Significant problems of ensuring sustainable development allow building decision trees to search for relevant frames. The attributes lying on the corresponding branch defining the frame do not guarantee that the full set of frames is taken into account. It is proposed to use decision trees to determine the importance of the basic features of sustainable development that define a certain frame, and then, based on the interdisciplinary area ontology, use the obtained coefficients for the entire ontology of the multiplex semantic network. To find the relevant frame, the meanings of the concepts that are most important for the corresponding frame are used.

Software has been developed that implements the proposed scientific and methodological approaches. Depending on the type of problem, the knowledge base stores: a list of frames and their corresponding solutions, the importance of adaptive ontology frames, types of relationships, and interdisciplinary research features.

4. Conclusion

Theoretical knowledge functions as a complex system of interdisciplinary interactions aimed at solving urgent problems of ensuring sustainable development. In the evolutionary process of the development of science, revolutionary changes occur in the schemes of cognitive activity, embodying a progressive system of ideals and norms of research. Methods of setting initial frames and links of interdisciplinary research based on data mining methods are considered. A list of factors for assessing the relevance of frames for interdisciplinary research has been determined. An algorithm for determining the coefficients of the importance of frames and links in the process of functioning of an intelligent decision support system by many criteria has been developed. Ontologies are used to define these criteria. For this purpose, new attributes (semantic constraints of the system of concepts and relations, the importance of concepts and relations) are added to the tuple defining the ontology (concepts, relations between concepts, interpretation of concepts and relations), which are used to calculate the necessary distances in the multisets metric space.

5. Acknowledgments
The article is prepared with the financial support of the Russian Science Foundation, project № 19-78-10055.

References

[1] Fonseca E, Caldeira E and Vilela P 2020 *Simulation Modelling Practice and Theory* **102** 102068
[2] Lazar N and Chithra K 2020 *Journal of Building Engineering* **32** 101450
[3] Bag S, Wood C and Kayikci Y 2019 *Resources, Conservation and Recycling* **153** 104559
[4] Wu B, Tian F and Zeng Y 2020 *Geography and Sustainability* **1** 25–32
[5] Sun L, Niu D and Xu X 2020 *Journal of Cleaner Production* **287** 125528
[6] Kroger M and Schafer M 2016 *Futures* **84** 64–81
[7] Eustachio J, Caldana A and Martinelli D 2019 *Journal of Cleaner Production* **241** 118383
[8] Gunnarsdottir I, Davidsdottir B and Sigurgeirsdottir S 2020 *Renewable and Sustainable Energy Reviews* **133** 110294
[9] Zangeneh P and McCabe B 2020 *Advanced Engineering Informatics* **46** 101164
[10] Pronichkin S, Chernyshova Y, Kholstov A and Mamai I 2021 *Journal of Physics: Conference Series* **2060** 012010
[11] Homar A and Cvelbar L 2021 *Ecological Economics* **183** 106950
[12] Geng X, Liang Y and Jiao L 2020 *Applied Soft Computing* **94** 106460
[13] Roobavannan M, Kandasamy J and Sivapalan M 2020 *Ecological Economics* **174** 106665
[14] Wu W, Wang W and Zhang M 2020 *Ecological Complexity* **44** 100873
[15] Yang L, Cormican K and Yu M 2019 *Computers in Industry* **111** 148–71
[16] Mkhinini M, Narsis O and Nicolle C 2020 *Computer Science Review* **35** 100223
[17] Machado F, Diniz A and Brandao L 2020 *Electric Power Systems Research* **191** 106907
[18] Bosch J, Alvarez C and Ayguade E 2020 *Parallel Computing* **97** 102664