Application of cognitive technologies for modeling of design and technological knowledge and automation of design activities

Galina Volkova¹*, Tatyana Turbeeva¹, Oleg Grigoriev², Alexandr Sobolev² and Ruslan Dikov³

¹MSUT “STANKIN”, 3a, Vadkovsky per., Moscow, 127994, Russia
²ISA RAS, 9, 60-letiya Oktyabrya pr., Moscow, 117312, Russia
³TU-Sofia, 1000, 8Kl. Ohridski Blvd., Sofia, Bulgaria

*cog-par@yandex.ru

Abstract. In this chapter the problem of modelling, representation and objectification of technical knowledge are considered. A meaningful analysis of scientific and technical publications and design and technological documentation has revealed the features of reflecting technical knowledge in them and presenting it in various forms. This allowed us to develop a model for marking semantic elements and their structures in scientific and technical publications and technical documentation based on the method of conceptual (semantic) modeling in the framework of the methodology for automating intellectual labor (MAIL). To extract problem-oriented knowledge from scientific publications and technical documentation, a model for marking semantic elements and their constructions was developed. The markup model defines the fixation of significant semantic elements and constructs that are explicitly expressed in the document structure. The formed documentation structure defined a step-by-step methodology for the formation and processing of electronic funds of scientific and technical documentation.

1. Introduction

The rapid growth of information technologies is currently due to the increased demand for information products and services, as well as the improvement of software and hardware to support them.

In the works [1,3], the evolution of intellectual labor automation technologies was noted as a change of technological paradigms based on the logic of the cognitive approach. At the same time, traditional programming (or "electronic data processing") corresponded to the monadic level of this approach. Structural programming (or modeling of structures for data and programs) corresponded to the structural level, and finally object-oriented and functional programming (or knowledge modeling) – to the contextual level of the cognitive approach.

And all three levels (contextual, structural and monadic) form one "layer" or "scale of complexity" when modeling the real world and "immersing" it in a computer environment.

It can be assumed that the logic of further development of information and technology paradigms will reflect the change of programming technologies at the same functional levels with a new "scale of complexity".
This is already evident in the technology of multi-agent systems, and in artificial intelligence technologies.

The "Renaissance" of interest in artificial intelligence technologies became possible due to an almost abrupt change in the spatial and temporal characteristics of the computing environment, namely, the amount of memory for storing data and the speed for processing it.

This is almost a "repetition of the past". A similar situation was observed when computers were introduced in design activities.

At the initial stage - the stage of task-based automation of project activities - only those sets of actions that were performed by a specialist in traditional processes (information, computational, empirical)

However, the development of formal methods for processing large amounts of data and optimizing them allowed us to switch to computer-aided design automation [1,2], which significantly increased the efficiency of this process.

The "cornerstone" of the translation of subject information and knowledge presented in natural language form and recorded in human memory and on traditional paper media, in a qualitatively different environment - environment of computer and network technologies and other formal language representation - is necessary to modeling a semantic "space / framework" that ensures the adequacy of all syntactic representations of subject tasks that are subject to automation in any areas of human activity.

The fundamental scientific problem is to developing a theoretical apparatus for multi-level representation of subject knowledge and for extracting this knowledge from documentary sources [3].

This problem is particularly acute in relation to the representation and objectification of technical knowledge.

2. Method for extract problem-oriented knowledge from scientific publications and technical documentation

The study of methods for modeling, processing and extracting information and knowledge from scientific and technical publications and design - technological documentation has allowed us to establish the following features and problems:

- lack of methodological justification for the proposed methods and their implicit specialization in problem areas;
- be preferential modeling of knowledge based on ontological technologies;
- insufficiency or lack of formal apparatus in the proposed methods;
- be identification of syntactic elements and constructions with semantic categories and their connections, etc. [6, 7, 8, 9, 10].

A meaningful analysis of scientific and technical publications and design and technological documentation has revealed the features of reflecting technical knowledge in them and presenting it in various forms [11,12,13,14,15,16,17,18,19].

This allowed us to develop a model for marking semantic elements and their structures in scientific and technical publications and technical documentation based on the method of conceptual (semantic) modeling in the framework of the methodology for automating intellectual labor (MAIL).

A distinctive feature of this method of semantic modeling is [2,3,4]:
- availability of two model representations (universal and subject);
- availability of multiple models with different levels of abstraction in each representations;
- availability of the presence of static and dynamic components and their links in each model;
- the existence of regularities for a formation of all component models only for a universal conceptual representation;
- a presence of restrictions imposed by the universal representation on the conceptual representation of subject problems;
- unambiguous identification of elements of static and dynamic model structures of any abstraction level and any representation.
To extract problem-oriented knowledge from scientific publications and technical documentation, a model for marking semantic elements and their constructions was developed.

The basis of this model is a detailed content analysis of technical documentation in the form of normative reference documents of various levels.

A detailed analysis of normative reference documents allowed us to identify and generalize the features of problem-oriented knowledge representation in scientific and technical documentation [5,20]:

- any document is either a method for obtaining /forming / selecting a technical solution, or a method for processing a technical solution in the form of various types of documents;
- any method is a complex structure for making / receiving decisions based on a variety of subject constraints or dependencies presented in various forms: text, table, graphic, combined, etc.

The problem of semantic modeling of scientific and technical texts is that means of a natural language the big variety of statements is represented:

- on the place of the problem being solved in the structure of higher-order issues;
- on the specific features and peculiarities of problem setting in a certain subject area (related to a certain activity);
- on the method of investigation and/or the method of solving the problem (finding means of solutions);
- on the peculiarities of the functioning of the means ensuring the achievement of a certain goal;
- the exact characteristic of the object function purpose, etc. [1,2,4].

Modeling a product (object or process) can be seen as forming/generating a plurality of its model representations at its lifecycle phases.

A model representation in any phase (or any process) can also be fixed as a product representation based on the i-th level of the process concerned (research, design, production, etc).

Formally, the model representation of the n-th product at the i-th level of detailing is as follows [5]:

\[ MP_i(n) = \langle Y^i, q(Y^i), X^i, q(X^i), R_{Y^i X^i} \rangle, \]

where

- \( Y^i \) - description of the process of product model formation at the i-th level,
- \( q(Y^i) \) - a description of process properties on the i-level,
- \( X^i \) - a description of the objects involved in the i-level process,
- \( q(X^i) \) - a description of properties of objects involved in the process at the i-level
- \( R_{Y^i X^i} \) - a description of the relationships between the process and the objects involved at the i-level.

So, for example, if we consider the design phase \( Y^i \), then \( i \) will be presented by the compound index and will reflect decomposition of design on stages, design procedures and design operations. \( X^i \) is a set of objects involved in the process of functioning of the created product at the i-th level of detailing of the design solutions.

The identified features were the basis for developing the classification of fragments on the sheets of any technical document.

The basis for the markup model is the division of the complete semantic presentation of scientific and technical publications and technical documentation into two parts:

- explicit, reflected in all forms of documentation information,
- implicit one that defines the whole set of knowledge that has been used in their writing.

The markup model defines the fixation of significant semantic elements and constructs that are explicitly expressed in the document structure.

The separated semantic elements include [2,4]:

- subject categories of three classes (Object, Characteristic, Value) and their instances,
- first kind subject dependencies of different degree of formalization and their instances.

The separated semantic structures include:

- binary relationships of subject categories of three classes and their instances, (if there is an explicit fixation);
- binary relationships of first kind subject dependencies of different degree of formalization and their instances;
- structures defining the connection of subject categories and first kind subject dependencies;
- constructions determining the refinement of subject categories by their instances (in the form of “areas of permissible values”);
- constructions determining refinement of first kind subject dependencies by their instances (if the texts reflect examples of problem solving).

The formed documentation structure defined a step-by-step methodology for the formation and processing of electronic funds of scientific and technical documentation:
- developing an electronic catalog for normative reference documentations;
- formation of a fund of electronic images of documents;
- fragmentation of electronic images of documents;
- recognition and/or editing of sets of fragments of electronic images of documents by type with subsequent marking of semantic elements and their constructions;
- developing conceptual models for the selected document, taking into account the identified limitations and fragment typology.

3. Results
The analysis and studies thus produced the following results:
- detailed analysis of technical documentation of the project organization, which allowed to establish and summarize peculiarities of presentation of problem-oriented knowledge in scientific and technical documentation;
- a structural model of technical documentation has been developed, including the following structural elements: fund, collection, document, sheet, fragment;
- a classification of fragments on technical documentation sheets has been developed and availability of reference connectivity of content elements at the level of fragments, documents, collections of documents has been established;
- a step-by-step method of formation and processing of electronic funds of scientific and technical documentation has been developed, which includes a process consisting of stages and procedures, rules of their implementation and forms of presentation of the result;
- formulated requirements for conceptual representation of problem-oriented knowledge;
- a model of the product in the form of a set of its model representations on phases of life cycle with different degree of detail is proposed and their formal description is carried out taking into account the system approach;
- the model of marking semantic elements and their structures has been developed to extract problem-oriented knowledge from scientific publications and technical documentation, which allows to record semantic elements and their structures, which are clearly expressed in the structure of the document.

4 Conclusion
1. A pilot version of the electronic catalogue support and enhancement tools has been developed.
2. A fund of electronic images of documents has been formed.
3. Developed an experimental version of means of support of automated fragmentation and formation of description of graphic fragments.

Acknowledgments
This article was prepared as part of project implementation No. 17-29-07056/19 “Development of models and methods of representation and processing of the problem-oriented knowledge taken from scientific and technical texts and design-technology documentation” under the RFBR grant.
References

[1] G.D. Volkova, Methodology for automating intellectual labor: prerequisites, basis, results, problems. Part 1, Intersectoral information service. 2, pp. 9-20, (2009).
[2] G.D. Volkova, Methodology for automating intellectual labor: prerequisites, basis, results, problems. Part 3, Intersectoral information service. 4, pp. 15-31, (2009).
[3] G.D. Volkova, Methodology for automating intellectual labor, p. 104, (2013).
[4] Volkova G. D., Conceptual modeling of project tasks: textbook. manual ,117p., (2016)
[5] Turbeeva T. B., Volkova G. D., Semyachkova E. G., Vinarskaya G. A., Kharitonova A. S., Abramov K. M., Gavrilov A. G., Conceptual modeling of design processes for complex engineering products ,Electronic journal: automation and control in mechanical engineering, No. 2(31), pp. 12-22, (2018)
[6] Alexandr Palagin, Sergey Kryvyi, Nikolay Petrenko Architectural and structural features of automation tools for extracting knowledge from natural language texts, International Journal "Information Technologies & Knowledge", Vol.7. Number 3, pp.261-271 (2013)
[7] Palagin A.V., Kryvyi S.L., Petrenko N.G. About automating the process of extracting knowledge from natural language texts, Natural and Artificial Intelligence Intern. Book Series. Intelligent Processing, №9,pp. 44–52 (2012)
[8] Ermakova L.M., Methods of extracting information from the text, Bulletin Of Perm University. Math Series. Mechanics. Computer science, Vol.1(9), pp.77-84 (2012)
[9] Alexander Nechiporenko, System for automated knowledge extraction from natural language texts, Proceedings of the international scientific and technical conference "Information systems and technologies-2003" (2003)
[10] Kuzin E. S., Concepts of information technology of functional-oriented design of applied information systems, Information technology, No. 1, (2000)
[11] Lavrenova O. A. Modeling of semantics of scientific and technical texts for AIS and its theoretical bases, Proceedings of the 8th all-Russian scientific conference "Electronic libraries: promising methods and technologies, electronic collections",(2006)
[12] O. P. Kuznetsov, V. S. Sukhoverkhov, Ontological approach to the assessment of the subject of scientific text , Design ontology,Vol. 6, No. 1(19). pp. 55-66 (2016)
[13] Babanov A. M., Two modern approaches to semantic modeling – ORM and ERMM, Bulletin of Tomsk State University, no. 3(28), pp. 46-56 (2014)
[14] Babanov A. M., Comparative analysis of semantic models used for designing database schemas Bulletin of Tomsk State University, №290, pp.251-260 (2006)
[15] Rawlinson D., Kowadlo G., Generating adaptive behaviour within a memory-prediction framework, PloS one, Vol. 7, no. 1, pp. e29264, (2012)
[16] Halpin T., Object-Role modeling: Principles and Benefits, International Journal of Information Systems Modeling and Design, no 1(1), pp.32-54, (2010)
[17] Sidorova E. A. Approach to modeling the process of extracting information from the text on the basis of ontologies, Ontology of design, Vol. 8, No. 1(27), pp. 134-151, (2018)
[18] Gavrilova T. A. Visual methods of working with knowledge: an attempt to review, Artificial intelligence and decision-making, no. 1, pp. 15-21, (2008)
[19] Valkman Yu.R., Tarasov V. B., From ontology of design to cognitive semantics, Ontology of design., Vol. 8, No. 1(27), pp. 8-34, (2018)
[20] A. Kutin, V. Dolgov, M. Sedykh, S. Ivashin, Integration of Different Computer-aided Systems in Product Designing and Process Planning on Digital Manufacturing. Procedia CIRP 67, pp. 476-481, (2018).