American-Eurasian J. Agric. & Environ. Sci., 15 (5): 824-831, 2015
ISSN 1818-6769

DOI: 10.5829/idosi.aejaes.2015.15.5.12639

Model of Documents Management for Academic and Research Universities on Basis Set Theory

[1]*M.N. Krasnyanskiy,* [1]*S.V. Karpushkin,* [1]*A.D. Obukhov and* [2]*A.V. Ostroukh*

[1]Tambov State Technical University, Department of "Computer-Integrated Systems in Mechanical Engineering", Tambov, Russian Federation
[2]State Technical University - MADI, Department of "Automated Control Systems", Moscow, Russian Federation

Abstract: We described the application of set theory to the construction of a mathematical model of electronic document management system (EDMS) of educational and scientific institutions, which are a special case of agricultural institutions. We studied the major problems of classical and electronic document management and analyzed the work of intellectual property protection department in the academic and research institutions. We developed a mathematical model based on interaction of sets of objects, users and operations as central elements of a document management system. We introduced the basic concepts and operations of a document management process, the interrelation of sets and considered the problems restricting access to objects. Based on the proposed mathematical model we developed and described the document management information system for the Department of intellectual property protection and described the structure of its database. The results of the scientific and practical studies can be used in the development of integrated document management systems for academic and research institutions including agricultural institute.

Key words: Electronic document management · Patenting · Set theory · Automation in education · Agricultural institute · IT-based management · Intellectual property protection

INTRODUCTION

The introduction of information systems of electronic document management (EDM) is necessitated by an ever-increasing volume of processed information, high reliability requirements to storage, processing speed and data transfer. An academic and research institution with multiple information flows connected to a single complex document management system is not an exception in the overall process of labor automation. These features are inherent for agricultural institutions as well as scientific and research centers in agricultural industry [1, 2, 3, 4]. The implementation of a large-scale, sophisticated EDMS is a laborious and time-consuming process, requiring a significant amount of resources. Construction of mathematical models for such complex systems is not a trivial task. Therefore, in this paper we will look at a separate subsystem of electronic document management limited to the scope of Department of intellectual property protection of an academic and research institution and the application of set theory to describe the subsystem software.

Subject Area Analysis: Any process of automation is preceded by a deep analysis of the subject area, processes occurring in the system, its constituent objects and subjects. In this case, we consider a system of document management for the Department of intellectual property protection of an academic and research institution in general has examined, but the processes occurred here are correspond to those in the departments of patenting in agricultural enterprises [5, 6]. As part of its work, the Department processes applications for registration of computer programs, databases and patents from academic and research institution employees and external users. Further, a package of documents required

Corresponding Author: M.N. Krasnyanskiy, Tambov State Technical University, Department of "Computer-Integrated Systems in Mechanical Engineering", Tambov, Russian Federation.

for obtaining a certificate of registration or patent is formed. After all the documents have been checked, the package is submitted to the Federal Institute of Industrial Property (FIPS). If the application is approved, the academic and research institution and the authors of the application are granted exclusive rights under the intellectual property law [7].

The subject area analysis identified several important problems to be solved using modern information technologies:

- Reducing time and costs of data processing and documentation;
- Replacing primitive local databases, Excel spreadsheets, text documents, etc;
- Incorporating fast and convenient feedback tools, as well as those permitting changes to the application without the authors' consent;
- Providing automatic mailing of certificates.

Statement of the Problem: Having considered the main processes and problems of paper document management in the Department of intellectual property protection, we set the tasks of the EDMS designing. These include designing of an electronic document management system for the Department of intellectual property protection, solving the problems of designing, processing and management of applications, arranging interaction between the authors and the Department staff members, automation of documentation [8, 9, 10]. To achieve these goals we propose to use Web-technologies for the development of the information system, as their application makes the implementation of 'client-server' architecture easier and faster and provides multi-user work with the system [11, 12]. The first stage in the development of the system is the description of a mathematical model based on set theory, which is the foundation of the EDMS.

Mathematical Model of the EDMS Based on Set Theory: In the given subject area the main object is a document which is subject to various operations resulting in its creation, changes or deletion. These actions are performed by a limited number of people, namely, users of the system. Each of these actions, depending on the type of tool and its current state, results in a change of certain intrinsic attributes [13, 14]. Thus, based on the above a document can be represented as a system with multiple objects, changing their attributes and states by the actions of multiple users:

$$S = < U, P >,$$

where S is a document management system; $U=\{ u_1, u_2,... u_N \}$ are set of objects (documents, projects, results of experiments, research and development (R & D), design projects (DP), etc.); $P=\{ p_1, p_2,..., p_Q \}$are multiple users [15];

$$u_i = < C_i, A_i, D_i>,$$

where u_i is the *i-th* object of document management; $C_i=\{ c_{i1}, c_{i2},... c_{iM} \}$ is a set of states of the *i-th* object; $A_i =\{ a_{i1}, a_{i2},... a_{iT} \}$ is a set of attributes of the *i-th* object with the corresponding elements of a set of attribute values $D_i=\{ d_{i1}, d_{i2},... d_{iT} \}$. The state of the object c_{im} is characterized by ordered pairs:

$$c_{im}=\{(a_{it}, d_{it}) \mid a_{it} \circledcirc A_i, d_{it} \circledcirc D_i, t=1..T_i \},$$

where T_i is the number of attributes of the object u_i. Then, to get the current state of the object u_i, you must apply the function of the state μ (u):

$$\mu(u_i) = c_{im}.$$

We also introduce the concept of a set of operations $O=\{ o_1, o_2, ... , o_L \}$. Each operation in changing the object u_i performed by the user p_q, can be represented by the function:

$$o_l(u_i, p_q): u_i \circledcirc u_i, c_{im} \circledcirc c_{ik}, (a_{it}, d_{it}) \circledcirc (a_{it}, d_{it}^*), \quad (1)$$

if the following condition holds

$p_q \circledcirc \{p_j \mid \varphi (p_x) \geq \varphi (u_i) . \varphi (o_l)\}$, (2) where $l=1..L$, $t=1..T_i$ and d_{it}^* are new values of a_{it} attributes of the object u_i after the operation o_l is performed and the object changes its state form c_{im} to c_{ik}. To perform the operation the user p_q must have the necessary level of access determined by the access function φ (x), where parameter x is the object of document management, user or operation and the result is a numerical value representing the specified user's access level or the desired value of access to the object or operation. Thus, if the user access level φ (p_x) is greater or equal to the product of object access levels φ (u_i) and performed operation φ (o_l), this operation can be performed by the user. Take for instance the read operation o_1 and editing operation $o_2$performed by two users p_1 and p_2, where, $\varphi (p_1) < \varphi (p_2)$ [16].

The read operation o_1 of the object u_1 can be performed by the users p_1 and p_2, if φ (p1)> φ (u_1) . φ (o_1) and φ (p_2)> φ (u1) . φ (o_1), respectively.

If φ (p_1) <φ (u_1) . φ (o_2) and φ (p_2)> φ (u_1) . φ (o_2), the editing operation o_2 of the object u_1 cannot be performed by the user p_1, but it can be performed by the user p_2.

Changing of the states will continue until the object reaches a certain achievable end state c_{im}, which means that the work with the object is completed at this point.

A unique feature of a document management system of an academic and research institution is that it stores not only conventional documents, but also projects, results of experiments, R&D, DP, etc. This feature entails the need to provide information not only about the current state of the object, but also its past states and, moreover, the possible states. The examples of possible states of an object include different versions of a scientific report, document or set of experimental results. In this case, the state tree of an object u_i is understood as a directed graph, where the elements of a set of states $C_i=\{c_{i1}, c_{i2},... c_{iM}\}$ are the vertices of the graph containing information on all states c_{im} of the u_i. A set of branches are elements of the set of operations O performed on the object by multiple users P [17, 18].

Thus, to construct a mathematical model given the specificity of an academic and research institution it is necessary to build the state tree C_i for each object u_i. Each vertex describes the state of the object c_{im}, resulted from the operation o_l performed by the user p_q on the object u_i, being in the state c_{ik}. The states of the object are described by ordered pairs of attributes and their values (a_{it}, d_{it}), which comprise the technical attributes of the object (size, date, type, etc.) and the values, variables, text and numeric fields, graphics, audio or any other type of information describing the object. Thus, the process of constructing a state tree involves the analysis of the object life cycle and its representation in the form of a directed graph, whose tree root is its original state c_{i1} of the object u_i and vertices of the graph are the states c_{im}. A simple state tree is shown in Fig. 1.

We assume that the object u_i is a document, the states c_{ij} (j = 1….4) correspond to different stages of editing (from the creation of document to its finalization). We use the following operations: o_1 is read operation, o_2 is editing operation, o_3 is save operation. Now, we consider two important operations on the state tree: the extension and compression of the tree. The tree extension is understood as the process of acquiring new states c_{ik} of the object u_i, resulting from the operations o_l. The tree compression is the process of deletion of states that do not satisfy the conditions or remain unclaimed by the users.

Let us consider some examples of the tree extension, where the document u_i is the object, whose state tree is shown in Fig. 1. The tree extension can be made by one user through document editing, which results in new versions of the document represented as states c_{i5}, c_{i6} (Fig. 2a). In the course of further work on the document user p_1 must come to a final version of the document, consisting of either one of the possible states c_{im}, or their union (Fig. 2b) through the operation of union o_U. When combining states, the object acquires the attribute values from different states, forming a new, unique object state (c_{i7}). The operation o_U is significantly different from the above considered operations o_l as several states $\{c_{ij}\}$ are involved in the creation of a new state c_{ik}. We represent this as the ratio:

$$o_U(u_i, p_q): u_i \rightarrow u_i, \{c_{ij}\} \rightarrow c_{ik}, (a_{it}, d_{it}) \rightarrow (a_{it}, d_{it}^{*}), \qquad (3)$$
$$c_{ij}=\{ c_{ij} \mid Y_U(c_{ij})>0, j=1..M \},$$

where Y_U is a set of conditions of bundling: $Y_U(c_{ij})>0$, if condition c_{ij} is valid for the bundling; $Y_U(c_{ij})<0$ otherwise.

Another example of the tree extension can be joint work on the same document (Fig. 3). In this case, operations are performed on the object by multiple users (p_1, p_2, p_3), each of which generates their own version of the document u_i. In this example, the state of the object should be reduced to a single state, which is the most appropriate (in Fig. 3 it is c_{i9}), however, there can be more than one resulting final state of the object.

Thus, after the final satisfying state of the object has been received, the user must determine which of the states of the object they want to save for future work and which of them have no value. Thus, the next stage is the tree compression.

Let us consider this process for both examples. In the first case, after the work on the document is completed, the author has 7 states of the object, including the initial and final. The state c_{i7} has the greatest practical value to the user. Suppose that the state c_{i3} also keeps important data (as one of the stages of the work), while the remaining states are to be deleted using the operation of compression of the object o_c. The operation o_c extends the existing set of operations O, but its implementation leads to deletion of states, rather than their change:

$$o_c(u_i, p_q): u_i \rightarrow u_i, C_i = C_i \setminus C_i^*, \qquad (4)$$
$$C_i^* \circledcirc C_i, c_{ij}^*=\{ c_{ij} \mid Y_{\tilde{n}}(c_{ij})<0, j=1..M \},$$

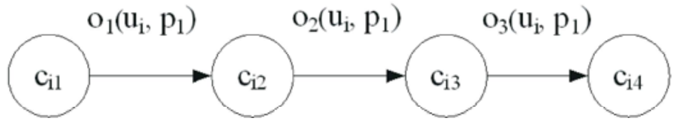

Fig. 1: A simple state tree of an object u_i

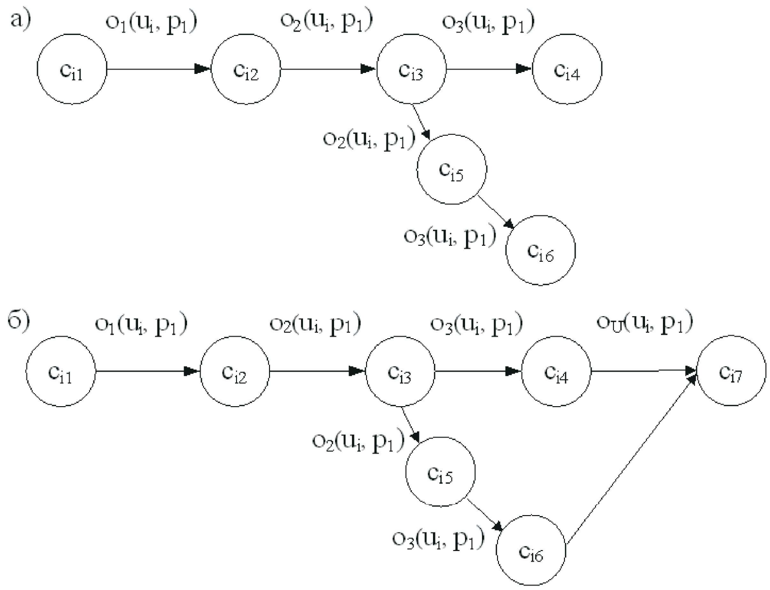

Fig. 2: State tree extension by one user

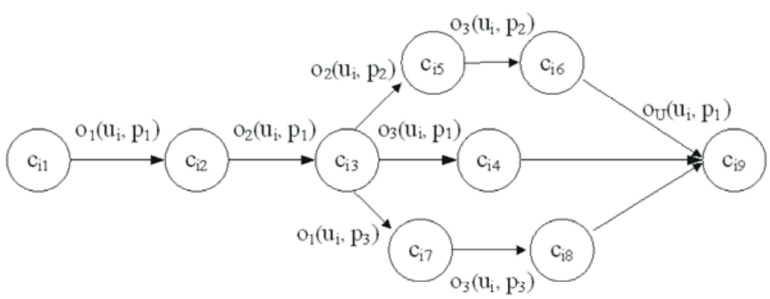

Fig. 3: State tree extension by multiple users

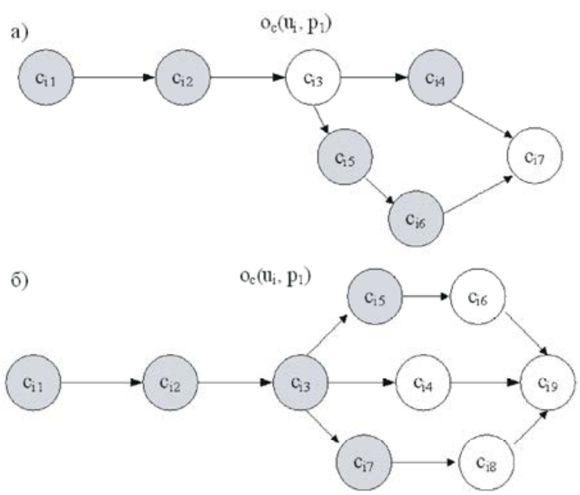

Fig. 4: State tree compression

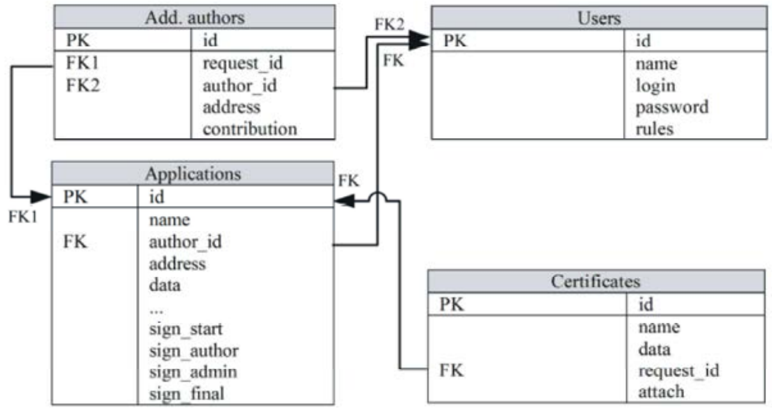

Fig. 5: A fragment of the database of document management information system

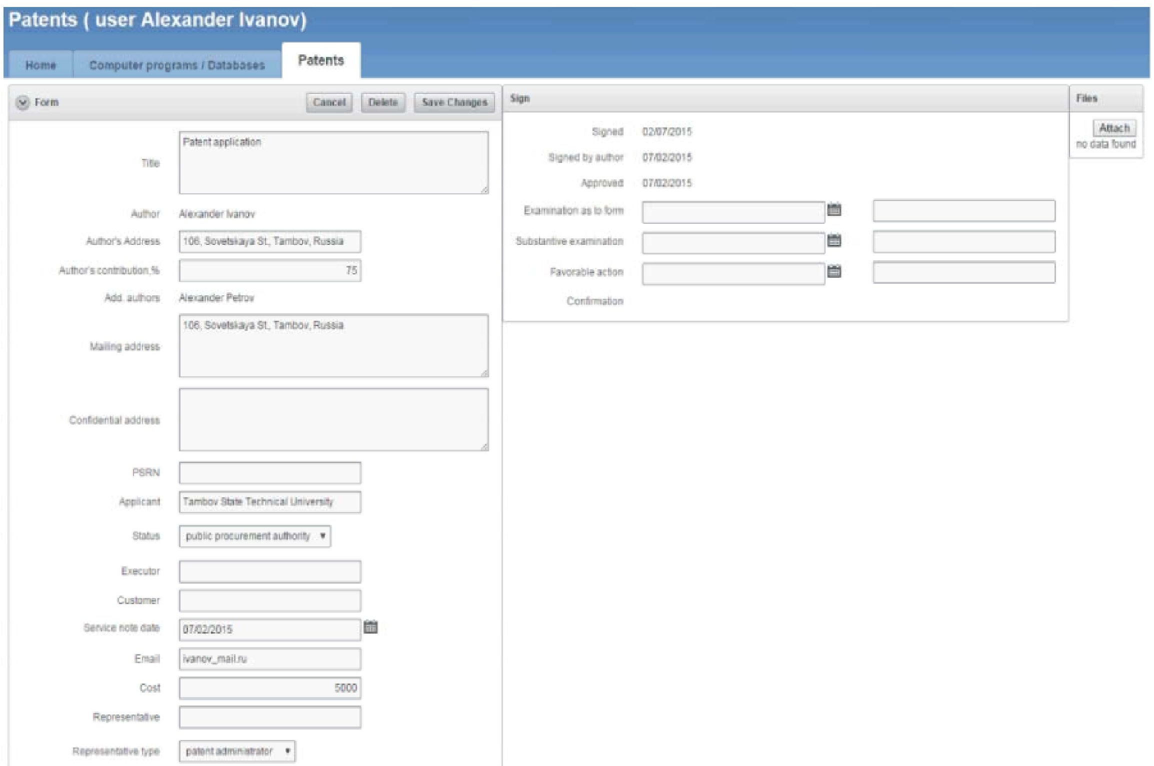

Fig. 6: Sample application form

where Y_c is a set of conditions of the value of the state: Yc $(c_{ij})> 0$ if c_{ij} state satisfies the conditions of the user; YC (cij) <0 otherwise. Thus, a subset of C_i * consists of the elements that do not satisfy the conditions of the user or the current task. After the operation of compression has been completed, the object stores only the states necessary for further work. We denote this by darkening the vertices of the remote states in the graph (Fig. 4a).

Thus, we examined the components of the mathematical model of document management between sets, basic attributes and operations on objects, the construction of the state tree of an object. Now, we turn to the main elements of designing the information system based on the given mathematical model.

Designing of the Information System: The resulting mathematical model is the basis for the designing of the database and the formation of the overall structure of the information system of document management, which we consider on the example of the Department of intellectual property protection of an academic and research institution [19, 20].

Multiple documents U = $\{u_1, u_2\}$ consist of the two basic forms of database tables – ‘Applications’ and ‘Certificates’. For each of them, we indicate entry boxes corresponding to the set of states and parameters. In our case, states are described by the entry boxes ‘application registered’, ‘checked by author’, ‘checked by administrator’, ‘completed’ and parameters are described by the fields ‘date’, ‘size’, ‘author’, etc. [21].

Elements of the set of users P are collected in the table containing data on logins, passwords and access rights of the user, on the basis of which the system determines which operations on the set of documents O the user can perform. Accordingly, the system of access to the object (2) has a simpler form: two types of users with high and normal levels of access, objects with similar access requirements and the two types of operations (public and administrative) [22, 23, 24].

Operations are implemented in the form of procedures, functions and queries to the database, which is fully consistent with the concept of operations in the developed mathematical model. The operations of bundling (3) and compression (4) are expressed in bundling and deletion of files, respectively.

As several users can refer to one document, this has led to the implementation of another table ‘Add. authors’, which describes the relationship ‘one-to-many’.

Figure 5 shows a fragment of the database underlying a document management information system for the Department of intellectual property protection of an academic and research institution.

Thus, we developed an information system in the form of a website with the following structure sections: ‘Computer programs / Databases’, ‘Patents’, all of which have sub-sections ‘Service notes’, ‘Applications’ and ‘Certificates’, corresponding to the main stages of document processing. First, the user enters the information on the in service note into the corresponding form and boxes and then, after it has been approved, gets access to the full application form (Fig. 6) [25, 26, 27].

Having entered all the necessary data and attached the accompanying files, the user sends them to the department staff member for processing. The staff member, in turn, blocks access to the author to edit and makes the required changes. If necessary, the author can withdraw the application for correction of errors. This arrangement provides fast data editing by both parties in real time.

After all the data have been approved, the staff member completes editing and downloads the application form and the document with the abstract formalized to the required standard on the computer. They are available both for printing and editing by means of Microsoft Office, if there is such a need. After the copyright for invention or computer program has been granted, the staff member uploads all the necessary certificates to the database and then, they are available to all authors and entered into their personal ranking [28, 29].

CONCLUSIONS

We described the process of designing the information system of electronic document management using set theory in the construction of the mathematical model. As the subject area we chose the Department of intellectual property protection of an academic and research institution. Chairs, which are similar in structure, have presented in agricultural institutions and plants and perform the task of patenting of inventions in agricultural technologies. The proposed approaches for development of automated control system of documents and building of information system are applicable for the processes and managing the documentation both in scientific and agricultural institutions.

We analyzed the subject area and created a mathematical model based on the interaction of multiple users, objects and operations. Operations act as a tool to move an object from one state to another. We introduced the concept of differentiation of access to subject using the function access that determines the possibility of performing an operation on the object by a particular user.

We considered the concept of state tree, the operations of extension and compression to describe the life cycle of an object. The first category includes operations that enable to generate new states of the object, while the second category includes a delete operation of states that do not satisfy these criteria.

Based on the proposed mathematical model we developed the document management information system for the Department of intellectual property protection. In this way, the results of conducted research will be used for describing and development of document management systems of scientific and educational institution, agricultural institute, scientific and research chair of agricultural enterprises.

ACKNOWLEDGMENTS

The work was financially supported by the Ministry of Education and Science of the Russian Federation as part of government target projects.

REFERENCES

1. Ostroukh, A.V., M.N. Krasnyanskiy, S.V. Karpushkin and A.D. Obukhov, 2014. Development of Automated Control System for University Research Projects // Middle-East Journal of Scientific Research, 20(12): 1780-1784.DOI:10.5829/idosi.mejsr.2014.20.12.21091.
2. Obukhov, A.D., M.N. Krasnyanskiy, S.V. Karpushkin, N.V. Molotkova, I.V. Galygina and A.V. Ostroukh, 2014. Electronic Document Management System Structure for University Research and Education // Journal of Engineering and Applied Sciences, 9: 82-189. DOI:10.3923/ jeas.2014.182.189
3. Albert Cervera, 2010. "Document Management in the Open University of Catalunya (UOC) Classrooms". D-Lib Magazine, 16 (7/8).
4. Duan, W.W. and D.M. Zhou, 2014. Study on Internet-Based Agricultural Production Informationization. Applied Mechanics and Materials, 608: 317-320.
5. Sarah Henderson, 2009. Personal document management strategies. Proceedings of the 10th International Conference NZ Chapter of the ACM's Special Interest Group on Human-Computer Interaction.
6. Ron Chimes, 2009. What Exactly IS Electronic Document Management? IDMi Information Document Management International.
7. Bernd Frohmann Revisiting, 2009. What is a document? Journal of Documentation, 65(2).
8. Jiang-Liang Hou, Fong-Hsin Lin, 2006. A hierarchical classification mechanism for organization document management. The International Journal of Advanced Manufacturing Technology, 28(3-4).
9. Kosuke Konishi, Naohiro Furukawa Hisashi Ikeda, 2007. Data model and architecture of a paper-digital document management system. Proceedings of the 2007 ACM symposium on Document engineering, 07.
10. Eero Eloranta, Ari Pekka Hameri, Mika Lahti, 2001. Improved project management through improved document management. Computers in Industry, 45(3).
11. Dingding Wang, Shenghuo Zhu, Tao Li, Yun Chi and Yihon Gong, 2008. Integrating Clustering and Multi-document Summarization to Improve Document Understanding. Proceedings of the 17th ACM conference on Information and Knowledge Management.
12. Sarah Henderson, 2005. Genre, task, topic and time: facets of personal digital document management. Proceedings of the 6th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: making CHI natural.
13. Stefan, Siersdorfer and Sergej Sizov, 2008. Meta methods for model sharing in personal information systems. ACM Transactions on Information Systems, 26(4): 1-35.
14. Krasnyanskiy, M.N., A.V. Ostroukh, S.V. Karpushkin, A.D. Obukhov, N.V. Molotkova and I.V. Galygina, 2014. Electronic Document Management Systems Structure for University Research and Education. Journal of Engineering and Applied Sciences, 9: 182-189.
15. Krukovskiy, M.Y., 2005. Grafovaya model' kompozitnogo dokumentooborota [Graph model of a composite document]. Matematichni mashini i sistemi, pp: 149-163.
16. Osborn, S., R. Sandhu and Q. Nunawer, 2000. Configuring Role-Based Access Control To Enforce Mandatory And Discretionary Access Control Policies. ACM Trans. Info. Syst. Security, 3: 2-21.
17. Novikov, D.A. and A.L. Sukhanov, 2005. Models and mechanisms for managing research projects in universities. M.: Institute of Education Management RAO, pp: 80.
18. Pechnikova, T.V. and A.V. Pechnikova, 1999. Practice working with documents in the organization: Textbook, manual for schools. M.: EMOS, pp: 208.
19. Banker, R.D. and R.J. Kauffman, 2004. The evolution of research on information systems: A fiftieth-year survey of the literature in Management Science. Management Science, 50: 281-298.
20. Obukhov, A.D. and M.N. Krasnyansky, 2014. Formirovanie dokumentov v Oracle Apex na osnove HTML shablonov [Formation of documents in Oracle Apex HTML-based templates]. Materialy mezhdunarodnoj nauchno-prakticheskoj konferencii "Virtual'noe modelirovanie, prototipirovanie i promyshlennyj dizajn" [Proc. Int.Conf. "Virtual modeling prototyping and Industrial Design"]. Tambov, pp: 326-331.
21. Ostroukh, A.V. and A.B. Nikolaev, 2013. Development of virtual laboratory experiments in iLABS environment. International Journal Of Applied And Fundamental Research, pp: 2.

22. Ostroukh, A.V. and A.B. Nikolaev, 2013. Development of virtual laboratory experiments in iLabs. International Journal of Online Engineering (IJOE), 9(6): 41-44. DOI: 10.3991/ijoe.v9i6.3176.
23. Ostroukh, A.V., KA. Barinov, A.B. Nikolaev and V. Yu. Stroganov, 2013. Interactive Game Modeling Concept for Personnel Training at the Industrial Enterprises. World Applied Sciences Journal (WASJ), 28(1): 44-55. DOI: 10.5829/idosi.wasj.2013.28.01.1876.
24. Ostroukh, A.V., K.A. Barinov and N.E. Surkova, 2013. Computer game modeling organizational structures of enterprises and industrial associations. Research Inventy: International Journal Of Engineering And Science, 3(12): 20-29.
25. Ostroukh, A.V. and N.E. Surkova, 2011. E-learning resources in vocational education. Saarbrucken, Germany: LAP LAMBERT Academic Publishing, pp: 184. ISBN 978-3-8433-2216-4.
26. Ostroukh, A.V., M.I. Ismoilov and A.V. Merkulov, 2012. Corporate training. Training of personnel of enterprises based on a virtual model of the professional community and grid technologies. Saarbrucken, Germany: LAP LAMBERT AcademicPublishing, pp: 129. ISBN 978-3-659-23865-9.
27. Ostroukh, A.V. and L.V. Vladimorov, 2012. Distance education technologies. Research and development of software products for video lectures and webinars. Saarbrucken, Germany: LAP LAMBERT Academic Publishing, pp: 97. ISBN 978-3-659-24981-5.
28. Ismoilov, M.I., A.B. Nikolaev and A.V. Ostroukh, 2013. Training and retraining of personnel and industrial enterprises of transport complexes using mobile technology. Saint-Louis, MO, USA: Publishing House Science and Innovation Center, pp: 166. ISBN 978-0-615-67111-6.
29. Ostroukh, A.V., K.A. Barinov, A.B. Nikolaev and V.Yu. Stroganov, 2014. Formal methods for the synthesis of the organizational structure of the management through the personnel recruitment at the industrial enterprises. J. Appl. Sci., 14(5): 474-481. DOI:10.3923/jas.2014.474.481.