

STRUCTURE AND FUNCTIONAL FEATURES OF DOCUMENT MANAGEMENT SYSTEMS ON THE EXAMPLE OF THE DEPARTMENT

Abduvalieva Zebiniso Abdulhamidovna¹, Marisheva Larisa Timofeevna¹ and Latipova Nodira Xalimovna¹, Sheyna Nataliya Evgenevna²

¹Tashkent University of Information Technology named after Muhammad al-Khwarizmi, Tashkent 100084, Uzbekistan

²Tashkent Technical University named after Islam Karimov

zebiniso_777@mail.ru, loramarish@mail.ru, nadira-latipova@mail.ru, evgeniya@mail.ru

Abstract. This article discusses the structure and functional features of document management systems in higher educational institutions. In the course of the work, the task of document management in the department is solved. Automatic recognition and optimization of documents provides advanced requirements management functions, such as automatic tracking, template compliance verification, guided editing and interaction between professors and students, requirements management. Documents in higher education institutions are often created in general-purpose formatted text editors, such as MS Word. The method given in the article contains instances of logical structures, such as document transfer, document flow optimization, a rule for checking graduation papers and a functional requirement. Differences in the content and physical representation of these instances create problems for their accurate recognition and extraction. To solve this problem, a structure is presented that allows 1) to specify logical structures in terms of their content, optimize and categorize them 2) to extract instances of such structures from documents with formatted text.

Keywords: Workflow Process, Uz Doc, Workflow Automation, Optimization, Workflow Algorithm

1 Introduction

Document management process — includes full or partial automation of document management, and compliance with procedural rules, transfer of documents in electronic form from one participant to another. The transfer of documents between the department and the dean's office, as well as between the dean's office and the credit department is an open access for teachers, the dean and the head of the credit department.

Open access includes the procedure for approving documents related to the educational process, managing the process of defending theses, managing meetings, and managing. In the works of Masuda Isaeva Technology Management Department Seoul National University Paperless university – how can we make it work? [1] a system for automating the learning process of students is given. The workflow process from document drafting to the implementation of the flow through the department, dean's office and other departments, such as signature, automation of the workflow process in an educational institution, consists in

workflow management, namely document management within the department itself and external sources, the components of which are the following elements: university teaching staff, resources, events, status management and also, the implementation and completion of the workflow, monitoring the implementation of the entire educational process [2]. The workflow process from document drafting to the implementation of the flow through the department, dean's office and other departments, such as signature, automation of the workflow process in an educational institution[3], consists in workflow management, namely document management within the department itself and external sources, the components of which are the following elements: university teaching staff, resources, events, status management and also, the implementation and completion of the workflow, monitoring the implementation of the entire educational process [4]. The main thing in document management is to simplify data transfer and make the workflow process easier and faster. In this paper, we are trying to find ways to solve the problem of document management using an information system. Please note that the first paragraph of a section or subsection is not indented[5]. The first paragraphs that follows a table, figure, equation etc. does not have an indent, either.

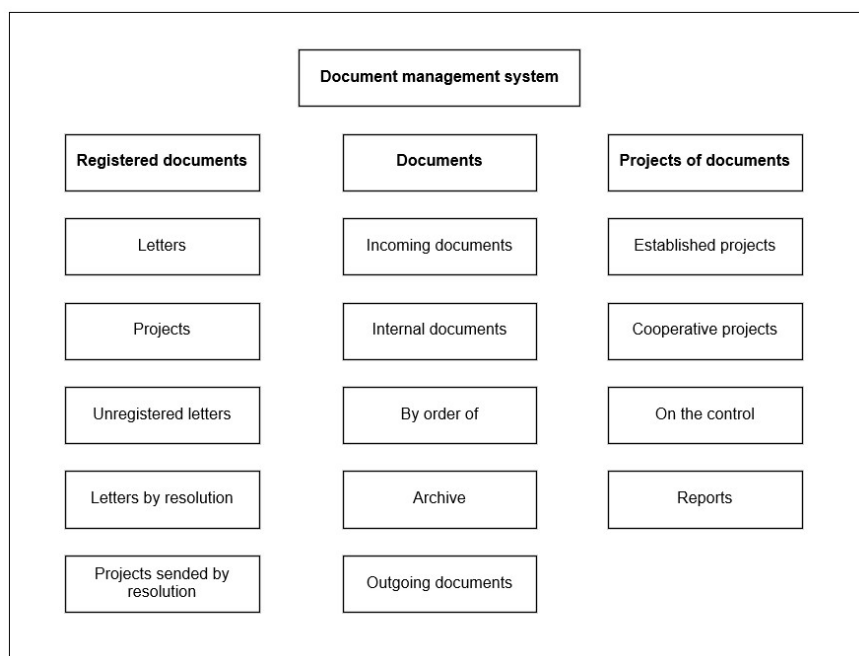


Fig. 1. Document management system

Subsequent paragraphs, however, are indented. Your submission may be prepared in LaTeX or Microsoft Word. Font will be Times New Roman. There is no need to include page numbers or running heads; this will be done at our end [6]

2 Materials and Methods

1) The UzDoc software tool has the following basic approach to implementation, based on the transmission and reception of data based on a common database; based on a combined mail model and a common database (i.e. group work mode). The implementation model is to transfer data from one person to another via email. Its characteristics are easy mode. However, the biggest disadvantage is the inability to monitor, no one can catch the dynamics of work processes. One of the advantages is the ease of data transfer [7]. Other problems include: difficulties with automated processing, such as archiving documents; data is easily copied; the transmission of many e-mail messages caused network traffic [8]. To overcome this disadvantage, automatic processing of the data warehouse, data transfer security, and ease of implementation are used.

The data processing and transmission model given in the UZdoc system is a model with lower development costs, easy to use, applicable to a wide range of document management tasks in a higher educational institution, system maintenance is extremely simple. The model encompasses the above advantages. The database manages information, and e-mail contributes to the workflow, as well as the possibility of joint development and work on projects [8]. Given the relatively high development costs, the implementation is relatively complex, the development and maintenance of the system is also of medium complexity. The developed model is considered suitable for a larger scale, collaboration does not require the use of high, difficult operations. During the development of the UZdoc system, document management was taken into account, the request in accordance with the functional requirements of higher educational institutions and the model of processing and transmitting documents, development costs, the scale of the system, the expansion of the field of activity of teachers, workflow automation systems in the department at the university. For the practical implementation of the above-described concept of optimizing electronic document flow, the following algorithm has been developed (Fig.2).

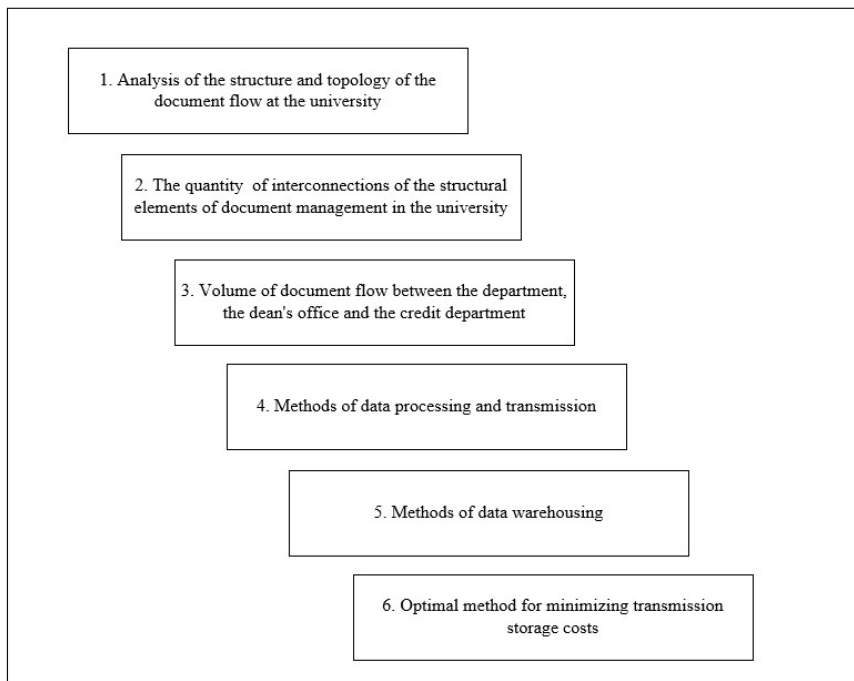


Fig. 2. Topology of the document flow management

2) As a result of the work , the following results were obtained:

- The algorithm of document flow is developed on the example of the department of "System and applied Programming",
- The Uz doc software tool has been developed, which allows the transfer and exchange of documents in the organization.

To determine the number of documents edited in a given period of time, you should have an accurate idea of which document is being received and what you need to get as a result[9]. The data will be needed to determine the number of current documents. How many are being created and how many will be edited.

Table 1. The direction of movement of documents.

Which department and to whom			
Incoming data	Who sent it	Outgoing data	Where to send

To determine the amount of data, whether it will be documents, letters or information related to the project, the volume of each of these data should be presented [10]. Any information has its own characteristics, in this case it can be digital, text, date, time, etc. Thus, let x_{ij} show that document j is stored in the i -th information store as follows:

$X_{ij} = 1$ if document j is stored in the i -th information storage;

0, if document j is not stored in the i -th information storage;

where

$i=1,2,3,\dots,n$ - the number of words in the document

$j=1,2,3,\dots,m$ - number of documents in the organization [11]

Suppose that each j -document can have a certain number of copies of r_j , then:

$$\sum_{i=1}^m x_{ij} = p_j \quad (1)$$

O_{ij} shows that j contains document i . Each document j has copies, denote them by k_r [12]

Next, we calculate the number of documents in a certain period of time.

$$\sum_{y=1}^m p_y * U_j \leq D_j \quad (2)$$

for $1 < i < n$,

where U_j -объем j th document, D_j - amount of available memory

p_y - periodicity of education, $y = 1, 2, 3, \dots, 12$

Having calculated the total amount of information that passes through the system during the year and the frequency of preparation of documents[13], we will get a representation of the total amount of information within one year

$$O_{q_v} = \sum O_{d_a} \quad (3)$$

for O_{d_v} - the total volume of all documents in the system.

To determine the amount of stored information, you should:

determine the characteristic properties (shelf life, storage need) [14]

- where the documents come from (via mail, or according to a certain information distribution scheme)
- archive construction structure
- editing the data storage method is automatic or manual. To do this, calculate the volume of stored documents

Graphically, this classification can be represented as follows:

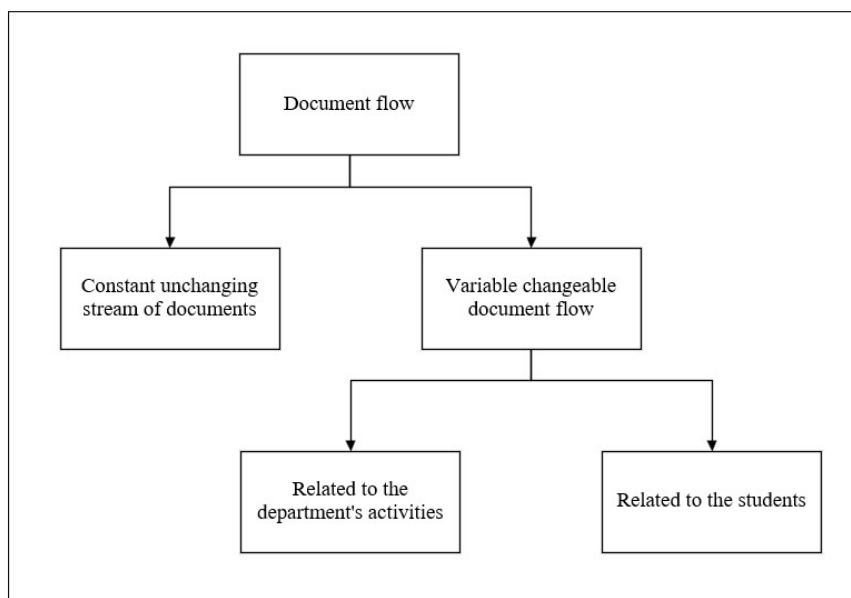


Fig. 3. Document flows management

3) Application Implementation. In the experiments, I used the documents of the Department of "System and Practical Programming" of the Tashkent University of Information Technologies. The dataset contains categories of department documents, including information about the department, about students, articles, etc. Each large class also contains several subcategories below and includes a total of 500 documents. In the experiment, the main categories are selected: department documents; documents related to teachers; and documents related to students, including the annual report, plans and rating. There are subcategories in each selected category. We have selected the main categories. For example, the annual report of the department means that we select the report for 2021 and 2022 as target categories, for example. The specific distribution of data is shown in Table 2.

The definition of the document flow structure consists in finding optimal ways of passing documents (information) within the organization according to a certain criterion. The solution to this problem is based on a comprehensive analysis of document flow models in the organization. The tasks arising in the analysis of complex systems require the use of methods of formalized description of each element of the system and their interaction. Depending on the formulation of the optimization problem, various approaches to modeling the structures of complex systems are currently used. The document flow model in higher educational institutions can be described using two types of structures - organizational and technical. The organizational structure describes the composition and order of interaction between the department, ensuring the solution of functional tasks. The technical structure describes the composition and interrelation of technical means of data storage and transmission [15].

In the process of work, it is often necessary to contact other organizations to interact and solve problems. In the example of a university, this may be cooperation with foreign universities to

work on projects or organize student exchange or cooperation. To do this, there is a section "cooperation" in the admin panel.

The department should also take into account how many employees work. Be able to add new teachers or staff. To do this, there is an employee management section and interaction with the HR department.

Table 2. Department data distribution.

Target dataset Initial	Initial training data	Auxiliary data
Students documents	Themes of diploma work and master's theses Written work by part-time students and reviews Documents on student practice	Documents on student practice
Plan	document control/documents to test students' knowledge (bases of written and oral control questions, written and electronic test questions, options for written work, block modules, and so on).	report on spiritual and educational work.
Scientific works	Reviews of scientific works	Diplomas, dissertations
Teachers documents	Documents of cooperation with professional colleges and enterprises of the department	articles of teachers of the department / Individual work plans of teachers of the department
Department documents	Orders of the dean of the faculty and information on their implementation / Decisions of the University Council, Methodological Council, Faculty Academic Council (copy)	Certificates and instructions of the rector and vice-rectors of the university on the activities of the department and their implementation

3 Results

The results after optimizing shown in table 3:

Table 3. The results after optimizing

Dataset	Before optimizing	After optimizing
Students documents	532	250

Plan	400	190
Scientific works	500	220
Teachers documents	380	230
Department documents	450	341

4 Conclusion and Further work

In this article, we proposed a critical path-based scheduling algorithm for workflow applications. To analyze performance, we compared our proposed scheduling algorithm with other existing heuristic scheduling algorithms for various sizes and types of workflow applications. Experimental results show that the proposed algorithm achieves certain goals and works better than these heuristic algorithms to reduce the cost. In the future, it will be possible to consider a more expanded structure of the functional features of document management in educational institutions, and we will be able to calculate the time of transmission of information and the quality of transmitted documents using various methods, such as dynamic calculation of the critical path. In addition, the proposed work can be expanded by using optimization algorithms of a different type.

References

1. Isaeva, M., Yoon, H., Y.: Paperless university — How we can make it work?. In: 15th International Conference on Information Technology Based Higher Education and Training (ITHET). pp. 1–8 (2016).
2. Luo, H., Fan, Y., Wu, C.: Overview of Workflow Technology. *J. Softw.* 11, 78-82
3. Fan, Yusun.: Base on Workflow Management Technology. Beijing:Tsinghua University Press, **32**, (2001)
4. Chen, Hong-na, Zu, Xu, Zhou, Feng: On the Developing Situation, Research Content and Trend of Workflow Technology. *Journal of Chongqing Instiute of Technology.* **20(2)**, 65-69 (2006)
5. Li, Zhao, Qing, Li, Farong, Zhong: A Visual Modeling Framework of Workflow Systems Based on CCS. Semantics, Knowledge and Grid. Fifth International Conference. pp. 200-207 (2009)
6. Dinesh, P, Mital, Goh, Week, Leng School of Electrical & Electronic Engineering Nanyang Technology University Nanyang Avenue, Singapur 2263. Text segmentation for automatic document processing. Rosemont, IL, USA. pp. 132-133 (1995)
7. Jian, LI, Na, YU. Research on Document Management System Based on Streaming Storage Technology. International Conference on Industrial Engineering and Engineering Management. pp. 558-562 (2011)
8. Ki Won Lee, Jeong Seok Kang and Hong Seong Park. DocT – Document Management and Testing Tool for Robot Software. International Conference on Ubiquitous Robots and Ambient Intelligence. IEEE. pp. 413-416 (2014)

9. Kruchinin S., V., Bagrova E. V., Chair of Economics, Management and Science Noyabrsk Institute of Oil and Gas (branch) TIU in Noyabrsk. Systems of Electronic Document Management in Russian Education. Pros and Cons. IEEE. pp. 628-630 (2019)
10. Li Sui, Gengchen Shi, Ping Song School of Aerospace Science and Engineering Beijing Institute of Technology Beijing, Xingyu Yuan Prosten Technology Holdings. Design and Implementation of ISO Document Management System. International Conference on Computer Science and Software Engineering. (2008)
11. Li Kai-Qi^{1,2} Diao Xing-Chun² Weng Nian-Feng^{1,2} Li Ge^{1,2}. Research on the Management Framework of Personal Desktop Document DataSpace. International Conference on Information Science and Engineering. pp 959-962 (2009)
12. Yong Wang Bi-yan Sun Fei Cheng Department of Information Engineering, Archives of WenZhou University R&D Center. Electronic-Document-Based Management Process Model for Image Archives in Universities. International Conference of Information Technology, Computer Engineering and Management Sciences. Pp 57-60 (2011)
13. Mirjana A. Andric, Member, IEEE and Wendy Hall. Using Metadata for Information Retrieval in Document Management Systems. Belgrade, Serbia. pp. 1093-1096 (2005)
14. asi Tyrväinen, Tero Päivärinta. On Rethinking Organizational Document Genres for Electronic Document Management. Hawaii International Conference on System Sciences. pp. 1-10 (1999)
15. Katarzyna Styk, Jakub Liszcz, Klaudia Drobek. Basic Project Management Documentation Based on the Example of the Student Project AGH Lean Line. pp. 45-49 (2019)