ALGORITHMS AND SOFTWARE FOR GENERATING DYNAMIC REPORTING IN AN ELECTRONIC DOCUMENT MANAGEMENT SYSTEM

Kurbanbaev Azamat Ikhtiyorovich
Student, "Computer Engineering" specialty, master's degree, Alfraganus University.
https://doi.org/10.5281/zenodo.10102487

Abstract. *Consideration of the issue that the implemented electronic document management system will satisfy absolutely all the requirements of the enterprise only if the system is properly integrated into business processes company and taking into account the peculiarities of internal corporate document flow.*

Therefore, for the full functioning of electronic document management systems, support and adaptation are required by the company implementing this or that system. The development of methods for integrating information resources is one of the most pressing problems in the field of information systems. It has begun to attract especially great attention in recent years due to the high heterogeneity of information data sources. The presence of various types of electronic documents makes the electronic document management system (EDMS) a highly heterogeneous information system.

Keywords: *programming, document processing, system, algorithmic, engineering*

The minimum amount of information in the EDMS is an electronic document. The main task of the EDMS is to store electronic documents, maintain a chronological record of their changes, ensure document flows throughout the enterprise, and control over processes that meet business requirements with which electronic documents are connected. With the introduction of an EDMS, a document encapsulates management tools (and actually becomes one itself) - all orders, instructions, decisions and others become a flow of documents.

There are two terms that are often confused: “paperwork” and “document flow”. Document flow is a definition used when declaring a specific set of actions when working with document metadata. There are so-called EDMS, to which you can set all the necessary document flow rules. There are systems for which the initial goal was to support various rules. Such systems do not have all the common features found in other systems that come close to being considered a fully automated EDMS. GOST on document flow and office work rules, which is adopted in Russia, is quite multifaceted, and creating a document flow system that satisfies all the provisions of the standard and rules is not a difficult and multi-criteria task.

The task of converting information from one form to another must be solved in every electronic document management system and, moreover, this task is the main task of such systems. The purpose of using electronic document management software systems is to reduce the volume of paper document flow (which is carried out entirely in human-readable form) by transferring some of the functions for exchanging documents to an automated system. Therefore, to achieve the goal, first of all, it is necessary to solve the problem of processing data into a form suitable for automated processing and back into a human-readable form. The processing of information from one form to another should be done automatically, that is, support for each specific format should not be associated with making changes and additions to the program code.

■ Storage requirements. If you have a lot of documents (from a storage perspective), you should choose a system that supports hierarchical structured storage (HSM - Bishop Storage Management). This mechanism allows you to store data that provides the fastest access. The most active data is not only the fastest data, but also the most expensive in storage media. At that time, the least frequently used data is automatically transmitted over slow and cheap means of communication.

The presence of pre-established rules that require the maintenance of a given artifact or type and the implementation of automated control. In particular, this includes preparing a report of a given sample and performing basic checks in the organization.

■ Presence in automatic document flow management with the organization. The degree of complexity of the organizational structure.

Presence of remote offices and divisions. This approach requires the use of more complex concepts for organizing data replication.

■ High-capacity archives. EDMS can be supplied with already integrated high-capacity form factors. Subsystems are installed with multiple drives and their own operating system.

■ The presence of a document flow system that does not meet current needs.

The requirement is to improve the management of document flow, as well as the management of the document life cycle and the process of its control. In response to this need, a proposal is put forward to control arbitrary business processes.

■ Shelf life requirements. With long storage periods, a significantly long period measured in more than one decade, the issue of replication arises, synchronous replacement of duplicates and equipment support.

■ Scalability requirements in terms of compatibility, system expandability. Possibility of integration with existing information systems and use of existing equipment. The need to store images of documents. Using the organization of certain document storage formats.

The need to support engineering and design tasks and other company features. The need for developed information retrieval tools. Full support for the language of the organization's document system.

■ Security requirements (encryption, access, etc.). The ability to use the organization’s existing infrastructure and information access mechanisms in the document management system. Requirements for compliance with certain standards: internal, industry, GOST, international quality control standards, level of organization of information storage.

To solve this problem of information processing, the author proposes and justifies the general mathematical model of information flows in the electronic document management software package he developed. The model consists of four main components (modules) - storage, loader (data loading module), saver (data upload module), checker (automated verification module).

The need to develop an internal format and dedicated data storage is justified by the following considerations:

• the electronic document management system provides different counterparties with the same information in different forms;

• within the system, it is therefore necessary to maintain meta-information about all forms of presentation of information circulating in the system;

• parts of this meta-information related to different forms of information presentation are necessarily different, and necessarily have quite a lot in common;

• external formats are controlled by counterparties and serve their current needs and requirements, and therefore may be subject to significant changes over time; The internal storage of the system should remain as stable as possible.

The data warehouse must provide the following functionality:

• the ability to quickly traverse the entire data structure and obtain arbitrary subsets of it;

• the ability to quickly save and restore data;

• transaction support;

• support for information download and bypass interfaces.

In order to ensure that the storage is filled with data coming into the system from counterparties, a loading module is needed, which is a buffer between the counterparties and the storage loading interface.

To manage metainformation, a set of auxiliary components combined in Fig. 5 logical block Schema. All components KDOM models process information using format descriptions, made in a specialized language.

This language is built on the basis of an extensible XML markup language, which is currently the industry standard for representing metadata. The meta-description language of the KDOM model allows you to describe syntactic rules for recognizing data elements in a machine-readable file, syntactic rules for generating machine-readable and human-readable data in accordance with the formats of external information consumers, as well as describe comprehensive rules for format and logical-arithmetic control of the validity of information in the system storage.

The most important properties of the KDOM system are its

• extensibility - the developer has the opportunity to create his own implementations of the Saver and Scanner components to connect the ready-made internal KDOM architecture with any external applications and user interfaces;

• flexibility - all descriptions of formats, checks, custom types, additional integrity constraints can be made in the form of XML schemas; the data type management mechanism allows you to describe domains of arbitrary structure and conduct automated checks of whether a property belongs to a domain;

• centralization - all data and metadata are consolidated in a single internal storage, which ensures speed and completeness of information processing;

• reliability - procedures for saving, restoring, integrity monitoring and checking external restrictions are used, providing intelligent control of the state of the internal storage, as well as a transaction mechanism that protects the storage from data structure violations during system operation.

Thus, the KDOM component model is a highly customizable software environment that can be used as a universal object model for the development of any processing and verification system information presented in the form of formalized documents.

The final section of the work contains a summary of the practice of implementing the electronic document management system designed by the author. Internal. The server architecture of the Kontur-Extern system is completely built on the KDOM model. Data storage in system memory during information processing organized as a chronological forest, and for long-term storage it is packaged in a DBMS as a corresponding relational table. Data exchange between

server systems and user terminals are built using the described protocol and document package format.

Electronic document management is based on the principles of business process management, during which documents are created and moved. Considering the fact that the concepts of electronic document management and electronic document management systems are inextricably linked, we can formulate several basic principles of electronic document flow, which to one degree or another have been highlighted by various authors:

-Providing the software platform used in a particular EDMS with the ability to store and search documents, as well as maintaining a messaging system between users.

- Provide support for distributed processing

- Information, i.e. the document should be sent only to those users to whom it is intended to create, approve and edit the document. This principle eliminates the possibility of erroneously sending a document, for example, to the wrong executor, as well as mass mailing of a document to all users instead of just one specific one.

- Scalability – determined by the set of supported platforms, the maximum number of users, and the number of levels of system structures. This principle determines the level of performance of the information system, i.e. its ability to work with the flow and volume of information that will be carried out after the implementation of the EDMS.

Openness of the architecture and the ability to integrate with other applications. The openness of the architecture is its flexibility; this principle is needed so that an organization implementing an electronic document management system has the opportunity to customize the system for a specific customer organization, removing some elements and adding others.
Integration is important, first of all, for ease of use of the system and speeding up the time of working with documents.

Thus, the goals of the dissertation work are achieved:

The constructed theoretical models and algorithms for converting and processing data in electronic document management software systems stand the test of practice and make it possible to create an application system that has no analogues in terms of document flow volumes served, scalability and flexibility.

- Support for various types of documents that the system works with. This principle involves working with documents that have different formats, as well as support for working with multiple versions document and the ability to track the relationship of documents.

- Collective work with documents. It consists in ensuring the possibility of collective work of a group of performers on one or more documents to better perform their execution task. This principle means that the system should not provide for rigid fixation of document execution routes. Execution of individual documents may differ from the standardly accepted procedure for executing a document of the same type in an organization, which may complicate the work on its execution and processing.

Monitoring the passage and execution of documents.

The head of the organization, or another person appointed by the controller for the execution of the document, must be able to fully and in real time monitor what stage of execution it is at or other document.

Identification of product customization features for needs specific customer. This principle is closely related to the principle of open architecture of electronic document management

systems. It lies in the ability to customize the system for a specific organization, taking into account the peculiarities of record keeping in it.

This principle is one of the most important, as it determines the very essence of the work of any EDMS, because storage, search and communication between users is the basis for the rapid creation of a document, its processing and storage.

LETTERS

The above principles of electronic document management allow for prompt and objective control over the passage of documents in the organization, and also quickly find out at what stage of processing a particular document is.

With documents of any content, it is necessary to perform certain actions: receive and send, register, transfer to their destination, monitor the progress of execution, impose resolutions. Specialized systems are precisely designed to solve such problems.

It should be noted that compliance with the above conditions and principles of electronic document management guarantees a better implementation of electronic document management systems. This happens due to the settings system that will take into account the requirements of a specific customer, the possibility of intermediate control over the movement of documents, as well as the requirement for special capabilities of the system to notify about violations in the process of working with documents and their passing. We cannot ignore the fact that the uninterrupted functioning of all office processes in the organization and their strict structuring is ensured by monitoring compliance document management systems to state standards and legislation in general.

Thus, by electronic document management we will understand this way of organizing office work in an organization, with in which the bulk of documents are created, used and stored electronically using a specialized class of software - electronic document management systems.
Electronic document management has a number of specific features that are associated with the above principles of electronic document management, as well as compliance with the necessary conditions for its use.

REFERENCES

1. Electronic document management systems [Electronic resource]. https://www.sites.google.com/site/upravlenieznaniami/tehnologii-upravlenia-znaniami/sistemyelektronnogo-dokumentooborota (accessed November 15, 2020)
2. DirectumRX in BP Russia: a unified information space for managing tasks and documents [Electronic resource]. Russia, 2018. Access mode: URL: https://www.osp.ru/resources/releases/?rid=38328 (access date: November 14.
3. Electronic document management system “E1 Euphrates”. Access mode: URL: https://evfrat.ru/ (access date 11/17/2020)
4. Requirements for the electronic document management system. Access mode: URL: https://www.atlassoft.ru/articles/trebovaniya-k-sisteme-elektronnogo-dokumentooborota (access date 11/13/2020)
5. Electronic document management system “DockVision”. Access mode: URL: https://docsvision.com/ (access date 11/18/2020)

6. Sabou Marta. Learning WEb Service Ontologies: an Automatic Extraction Method and its evaluation. //URL: http://kmi.open.ac.uk/people/marta/
7. Fraunhofer Institute Fokus http://www.sweb-project.org.
8. Sibley E.H., Taylor R.W. A data definition and mapping language. Comm. Of the ACM, v. 16, 1973, pp.750-759.
9. Sibley E.H., W.T. Hardgrave, Kogalovsky M.R., Makalsky K.I. A conceptual model to support multi-model external views. Proc. of the Joint U.S.-U.S.S.R. Seminar «Data Models and Database Systems». Austin, Texas. Texas University, 1979, pp. 146-187.
10. Smith K. Michael, Welty Chris, McGuinness L.Deborah. OWL Web Ontology Language. Guide. //URL: http://www.w3.org/TR/owl-guide/
11. Swartz A. The Semantic Web In Breadth // URL: http://logicerror.com/semanticWeb-long
12. Thieme C., Siebes A. Scheme Integration in the Object-Oriented Databases. CAiSE 1993, pp. 5470.
13. Tobies S. Complexity Results and Practical Algorithms for Logics in Knowledge Representation: Ph.D. Dissertation. Aachen, Germany, 2002.
14. Ullman J. D. Information Integration Using Logical Views // Proc. of the 6th International Conference on Database Theory (ICDT'97): Lecture Notes in Computer Science. Berlin. Germany: Springer, 1997. Vol. 1186. P. 19-40.
15. Vacura M., Svatek V., Smrz P. A Pattern-Based Framework for Uncertainty Representation in OntologiesText, Speech and Dialogue. Proceedings of the 11th International Conference TSD 2008, Brno, Czech Republic, September 8-12, 2008. / Eds. P. Sojka, A. Horak et al. LNAI 5246. Springer-Verlag, 2008.
16. Van Heijst, G., Schreiber, A. T., and Wielinga, B. J. 1996. Using Explicit Ontologies in KBS Development. //URL: http://ksi.cpsc.ucalgary.ca/KAW/
17. Wen-Syan Li, Chris Clifton. Semantic Integration in Heterogeneous Databases Using Neural Networks. Int. Conf. on VLDB, 1994, pp. 1-12.
18. Wiederhold G. Mediators in the Architecture of Future Information Systems. IEEE Computer 25:3, pp. 38-49, 1992.
19. Wielinga, B. J. and Schreiber, A. T. Reusable and sharable knowledge bases: a European perspective // Proceedings of First International Conference on Building and Sharing of Very Large-Scaled Knowledge Bases. Tokyo, Japan Information Processing Development Center. 1993
20. XQuery 1.0: An XML Query Language. W3C Working Draft, 15 November, 2002.