Standardization and technical regulation involve the digitization of international and regional standards into "smart" standards, introducing new requirements for the presentation of complex industrial manuals. The introduction of artificial intelligence and machine learning into the process of standardization, improving process automation, data analysis will provide an opportunity to create innovative standards. Creation of machine-readable standards and their presentation in the form of a database will contribute to better interoperability between systems.

The article demonstrates the method of ontological engineering for the automation of partial methodological examination of documents in accordance with the National Standardization Plan.

As part of the experiment, an independent tool was developed, based on the extraction of qualitative data, using an uncontrolled automatic keyword extraction algorithm. The algorithm does not require the creation of a learning corpus, can be applied to any text and language, and does not have limitations on the amount of processed data. The algorithm provides adjustment of the parameters of the repetition of keywords, with the possibility of taking into account less important keywords.

The work presents the following components: model classification standards, essence extraction module, categorization, thesaurus formation, monitoring and analysis. Calculations were carried out using formalization, determination of the target function, establishment of similarity measures, and description of document features. The work presents the theoretical aspect of application, design and description of the functional system

Keywords: smart-standard, standard development, identifying keywords, harmonized standard, national standard

UDC 004.822

DOI: 10.15587/1729-4061.2023.288542

DEVELOPMENT SYSTEM FOR COORDINATION OF ACTIVITIES OF EXPERTS IN THE FORMATION OF MACHINESCHETABLE STANDARDS IN THE FIELD OF MILITARY AND SPACE ACTIVITIES BASED ON ONTOLOGICAL ENGINEERING: A CASE STUDY

Anar Utegenova
PhD
Department IT Engineering
Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev
Baytursynuli str., 126/1, Almaty, Republic of Kazakhstan, 050013
Akylbek Bapyshev
Master of Science*
Zhanna Suimenbayeva
Master of Science*
Alisher Aden
Master of Science*
Tansaule Serikov
Associate Professor
Department of Radio Engineering, Electronics and Telecommunications
S. Seifullin Kazakh Agro Technical Research University
Zhenis ave., 65, Astana, Republic of Kazakhstan, 010011
Ruslan Kassym
Corresponding author
PhD, Senior Lector
Department Information Communication Technologies
Academy of Logistic and Transport
Shevchenko str., 97, Almaty, Republic of Kazakhstan, 050013
E-mail: kasym.ruslan@gmail.com
*INT-SAT Alatau LLP
Abay ave., 20/5-18, Almaty, Republic of Kazakhstan, 050013

Received date 07.08.2023
Accepted date 13.10.2023
Published date 30.10.2023

***How to Cite**: Utegenova, A., Bapyshev, A., Suimenbayeva, Z., Aden, A., Serikov, T., Kassym, R. (2023). Development system for coordination of activities of experts in the formation of machineschetable standards in the field of military and space activities based on ontological engineering: a case study. Eastern-European Journal of Enterprise Technologies, 5 (2 (125)), 67–77. doi: https://doi.org/10.15587/1729-4061.2023.288542*

1. Introduction

The modern development of all industries is acquiring the most important status in the context of a rapidly changing world and digital transformation. Strengthening the Republic of Kazakhstan's defense industry and preparing for joint ventures with other countries may provide an opportunity to implement reform in the areas of the military-industrial complex. The relevance of the topic is indicated by factors such as increasing demand for defense industrial complexes, special attention should be paid to compliance with regulatory requirements and legal requirements. component in this industry. Despite the openness of international standards [1], it is also necessary to take into account the requirements of existing standards in Republic of Kazakhstan. For example, the national database of normative documents on standardization in the field of space activities is more than 100 national standards. The main aspect of the problem raised is the unnaturally small

number of standards that do not meet the growing for the transfer of scientific and technical information [2], equalizing the quality of products [3] manufactured in different countries [4], ensuring the interchangeability of complex products [5, 6], as well as promoting international relations.

The study is considered controversial not only in Republic of Kazakhstan, but also abroad, the implementation of the construction and commissioning of industrial facilities should take into account part of the military-industrial complexes, which have dual use, therefore, when developing standards, both military and civilian standards should be taken into account [2].

New research is required in the development of Smart Standards (Standards Machine Applicable, Readable and Transferable) [3–5] for the field of military and space activities, which will make it possible to change the visualization of the standard [6] as a full-fledged electronic guide to optimize the workflow, provide an opportunity to get acquainted with the chronology of the creation of the national standard, the transparency of all customers, participants, performers and experts. A particularly important aspect is the development of additional tools to ensure the uniformity of the terminology of national military standards in official languages.

Of the proposed review, only a few studies were able to indirectly touch upon the development of an administration system for managers when updating and harmonizing national standards. In addition, not many studies are devoted to the transformation of knowledge from the usual type of standards into a machine-understandable form. Consequently, studies on the exchange of experience and diligent transfer of existing technological solutions and European developments, the interaction of the working groups of International Organization for Standardization (ISO), International Electrotechnical Commission (IEC), European Committee for Standardization (CEN), European Committee for Electrotechnical Standardization (CENELEC) and other major standardization organizations will make it possible to adopt a single classification of formats for the presentation of standardization documents.

All these factors prompt the modernization, first of all, of the mandatory requirements system itself, as well as the development of an intelligent control system in the processes of developing standards for developers.

2. Literature review and problem statement

A review of existing approaches showed the multiplicity and multidimensionality of research trajectories. A thorough study of the problems presented in the research results in the development of key technologies in the management of civil-military integration standards [7] includes the study of various scientific approaches and technologies used to analyze the processes of development and harmonization of national standards. One of the ways to overcome these difficulties may be research [8], which presents convincing and influential arguments in favor of alignment standards and technologies, taking into account asymmetry, in identifying optimal harmonization strategies for states with an average level of technological development. But the problems of terminological unity, which directly affect the quality of the harmonized standard, remained unresolved. A number of scientific papers [9] raise the issue of introducing a criterion for measuring the level of harmonization of standards, which will help to more accurately determine how consistent the current national standards are with international requirements and expectations. However, when introducing machine-readable standards, these criteria are subject to careful analysis and revision by the industry of application, since the issues of creating different templates of Smart standards depending on the industry are widely discussed.

Particular attention should be paid to works devoted to the recognition of textual information, such as research [10] based on the identification of key terms for the industry standard that are related to legal categories, but questions remain about training the system and the time spent on training. After the stage of key word selection, metrological certification of the algorithm for assessing the inconsistency of these key comparisons of national standards is necessary [11].

In connection with the growing requirements for the training of industry specialists, machine-readable standards could act as educational and methodological complexes [12, 13]. However, questions arise regarding quality control and compliance of existing educational programs with current national standards. If this criterion is met, we would receive educational programs that meet industrial and international requirements [14–16]. But the problems of using these criteria are not covered and do not have practical examples. Technological techniques for extracting keywords are of scientific interest to the world community, comparison of algorithms provokes new questions for solving the problems of preparing and producing machine-readable standards [17, 18]. Most of the presented studies are focused on extracting qualitative data from natural language text, although this is an important aspect in the formation of machine-readable standards, however, great difficulties may arise when extracting data from graphs, digital models, diagrams and tables [19, 20]. The presented questions are also an indispensable resource for the extraction and subsequent processing of qualitative data. And as a result, it is necessary to combine and correlate data elements to establish a relationship and supplement the available resources using semantic technologies. Accumulation of a store of knowledge: semantic analysis of standards, targeted extraction of content and its use in subsequent processes are reflected in the following studies [21, 22]. The current critical turn presented in studies of the applicability of the use of the presented practices will emphasize the role of fast and reliable synchronization of normative reference data with the help of automated integration processes, especially in relation to partial and heterogeneous changes in terminology and taxonomies. But the problems of structuring and presentation of information in dictionaries remained unsolved.

All the studies presented indicate that it is advisable to conduct research on the topic of coordinating the actions of experts when forming the format of machine-readable standards in the field of military and space activities based on ontological engineering. In the new industrial age, a standard is no longer solely a reference for engineers and technologists; it turns into a real model or database that can be integrated into information systems, software solutions and digital production lines. This approach allows us to develop innovative and competitive products that are in demand in global markets.

3. The aim and objectives of the study

The aim of the study is: to develop an intelligent automated system for managing, monitoring, and controlling the processes of the operational cycle of standard development using technological integration of optimal methods for creating a single metadata repository.

To achieve this aim, the following tasks are solved:

– the task of implementing a procedure for automatically assigning an updated or harmonized standard to one or more classes;

– develop a software module to solve the problem of harmonization of international standards, classification of standards, extraction of key terms, categorization, thesaurus generation, monitoring and analysis of development stages.

4. Materials and methods

The object of the study is methods of transformation of knowledge from the usual type of national standards into a machine-understandable based on onotological engineering.

Research hypothesis: If it is determined what methods and technologies, algorithms we need to develop for an adaptive, commutative-cognitive platform for developers of national standards, taking into account information security, then we will be able to create a working basis for the full-fledged work of the Republican State Enterprise "Kazakhstan Institute of Standardization and Certification". Therefore, we will be able to automate the process of communication between interested bodies and developers of national standards, which will be followed by a quick response to all technological needs of the Republic of Kazakhstan.

The study made the following assumption: Machine-readable standards should be based on a certain set of individual characteristics of the region, existing regulatory requirements and global trends. If this basis can be formalized and structured, as well as documented in complex industrial manuals, such information will become the starting point for the development of new approaches to the design of industrial complexes that will complement each other, since a full-fledged manual in the form of a Smart standard will outline the entire life product cycle, up to disposal.

The simplifications adopted in the work are that at the moment, for the production of one product, sometimes it is necessary to use more than 10 international standards that require deep harmonization. This is not the only problem for established industries, however, for newly emerging enterprises there are problems of lagging behind, ranging from technical modernization, harmonization of regulatory requirements, international experience, international standards, environmental monitoring requirements and, ultimately, disposal methods. The technical result of the proposed solution is to simplify the process of creating a flexible environment for complex industrial management in the form of a Smart standard and expand the range of functionality for managers when designing it.

Before moving on to the consideration of general methods, it is necessary to summarize and highlight the key problems of the research topic. The study includes key aspects when implementing smart standards:

1. Organization of terminological accuracy.
2. Coordination of presentation form: text containers. graphical and numerical data, and digital models.
3. Identification of requirements for the form of representation of semantic units (description, attributes, meanings).
4. Form of presentation of digital modeling, model ontology, and its components.
5. Determination of the product life cycle.
6. Reuse of knowledge.
7. Correlation with external regulatory requirements.
8. Development of a data integration system.
9. Development of a management system, expert-analytical system.

Each of the highlighted aspects involves the use of existing technological solutions using new approaches and methods.

As part of the research objectives, the ontological engineering methodology is used to automate the processes of harmonization of standards according to the National Standardization Plan, using the results of the analysis of various models and methods, including metadata, digital footprint, automation of text markup and intelligent management in standards development processes.

At the same time, the ontological representation of standards performs the task of classifying them into sections. For each document, its belonging to a specific section is determined. Classification is carried out at two levels: the first level is associated with classification groups, and the second level determines whether the document belongs to standards that are close in content.

The degree to which a document belongs to a particular category is determined based on the coincidence of keywords. The more coincidences, the higher the degree of belonging. This takes into account the possibility of relating one document to several categories.

The system also provides expert group leaders with a unique opportunity to see applications for harmonization of standards that correspond to their profile, taking into account information, keywords, and expert publications.

The task presented in the article implements the procedure for automatic assignment of an updated or harmonized standard to one or more classes. Each standard that enters the system must be processed by marking up the text and compiling a set of keywords that characterize this standard. Each standard, having its own set of key terms, is subject to automatic processing to assign it to a group of related standards.

Solution methods:

– For each pair <document, category> the degree to which the document belongs to the category, and this degree is a number ranging from 0 to 1 (the higher the degree of affiliation, the more the document fits this category);

– Set-theoretic relations can be defined on a set of categories. For example, multiple documents that make up a category may or may not overlap, i.e., the same document may belong to more than one category. Composite categories can be defined – a category can be a subset of another category, and so on.

Solution methods. Formalization:

– multiple categories (classes);

– a set of documents;

– an unknown objective function is a display that classifies a document and gives a degree of ownership.

It is necessary to define this objective function.

Solution methods. Setting the Objective Function:

– the most popular way is to set a similarity measure on multiple documents. For each class, a centroid document is specified, and the function value is a measure of similarity between the search image of the document and the image of the centroid of the class.

Feature space:

– definition of features – formation of a document search profile (DSP) – a vector of characteristic features of a document, used in the future to make decisions on working with the document;

– the DSP is a multidimensional vector in the space of document features that characterize the semantic content of the source document;

– a feature is a $f:P \rightarrow D_f$ display where D_f is the set of valid values of the characteristic;

– if features $f1,...,f_n$ are specified, the vector $d=(f1(d),..., f_n(d)$ is called the feature description of the document d∈P;

– feature descriptions can be identified with the documents themselves;

– in this case, the set $P=Df1\times...\times Df_n$ is called the feature space.

Types of features:

– binary characteristic: $D_f=\{0, 1\}$ (standards organization);

– ranked attribute: $D_f \in [0,1]$ (measure of proximity);

– nominal characteristic: D_f – finite set (family of standards);

– ordinal characteristic: D_f – finite ordered set (branch);

– quantitative attribute: D_f – set of real numbers (year of publication);

– hierarchical characteristic: D_f is a finite partially ordered set (terms in the thesaurus).

There are 4 types of features in the document that represent search characteristics:

– basic;

– main;

– additional;

– complex (computable).

Schemes for the presentation of search features of a document:

– basic features are present in the description of each document;

– depending on the document type, different characteristic extension schemes can be used;

– currently, there are about 300 search signs. Search features are divided into two types: primary and secondary.

Fig. 1 shows the BPMN (Business Process Model and Notation) diagram of the process. Business process modeling allows to find a compromise between a clear description and technical functionality.

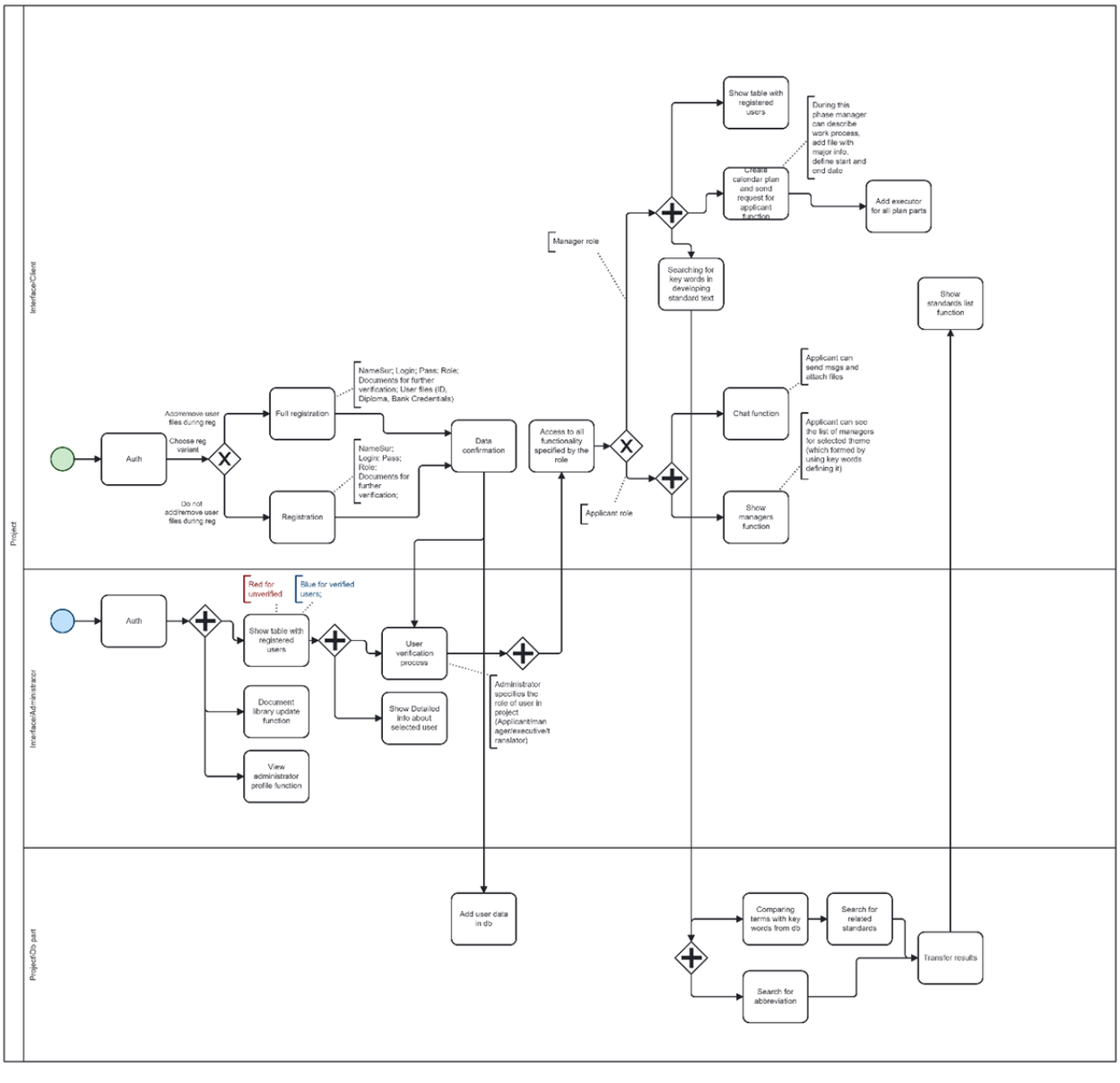

Fig. 1. Business Process Model and Notation process diagram

To implement the system development process, the following Toolkit was used:

1. Tiered Compilation in B.NET. In simple terms, multi-level compilation reduces application startup time.

2. C# programming language. This programming language was chosen because of its ease and accessibility. C# is a modern object-oriented and safe programming language. C# is part of the well-known C family of languages, and will feel familiar to anyone who has worked with C, C++, Java, or JavaScript.

3. PostgreSQL 15 contains many new features and improvements, including: Support for the MERGE SQL command. Selectively publish Table contents as part of logical replication with the ability to specify column lists and row filters. More options for using compression and support for Zstandard (zstd) compression. Including the ability to compress on the server side when creating a copy using pg_basebackup. Support for structured server log output in JSON format. Performance improvements, such as optimizations for in-memory and on-disk sort operations.

4. Visual Studio. An information system requires the use of a database. To work, it is necessary to store methodological works directly in the database. Visual Studio has a local database that is very convenient to use. All query syntaxes are common among all database languages and can be easily adapted. MySQL is a free database where the flexibility of the MySQL DBMS is expressed by supporting a large number of table types.

As a result, this paper presents: a model of classification of standards, a module for selecting entities, a categorization module, a module for the formation of a thesaurus, a module for monitoring and analysis. The calculations were carried out using the methods of formalization, determination of the objective function, similarity measure, as well as a feature description of the document. The results were tested by expert specialists with experience in the military and aerospace industries.

5. Results of research of development of a system of management, monitoring and control of development standards

5. 1. Description of the database structure and system functionality

A database model describes the various tables and their relationships in the context of a project. The important components of the model are presented in Table 1.

In UML (Unified Modeling Language) terminology, a domain conceptual model is essentially a class diagram (Fig. 2). Typically, this model omits most of the details, such as the attributes and operations of the classes, because the purpose of the model is to assign responsibilities to the entity classes.

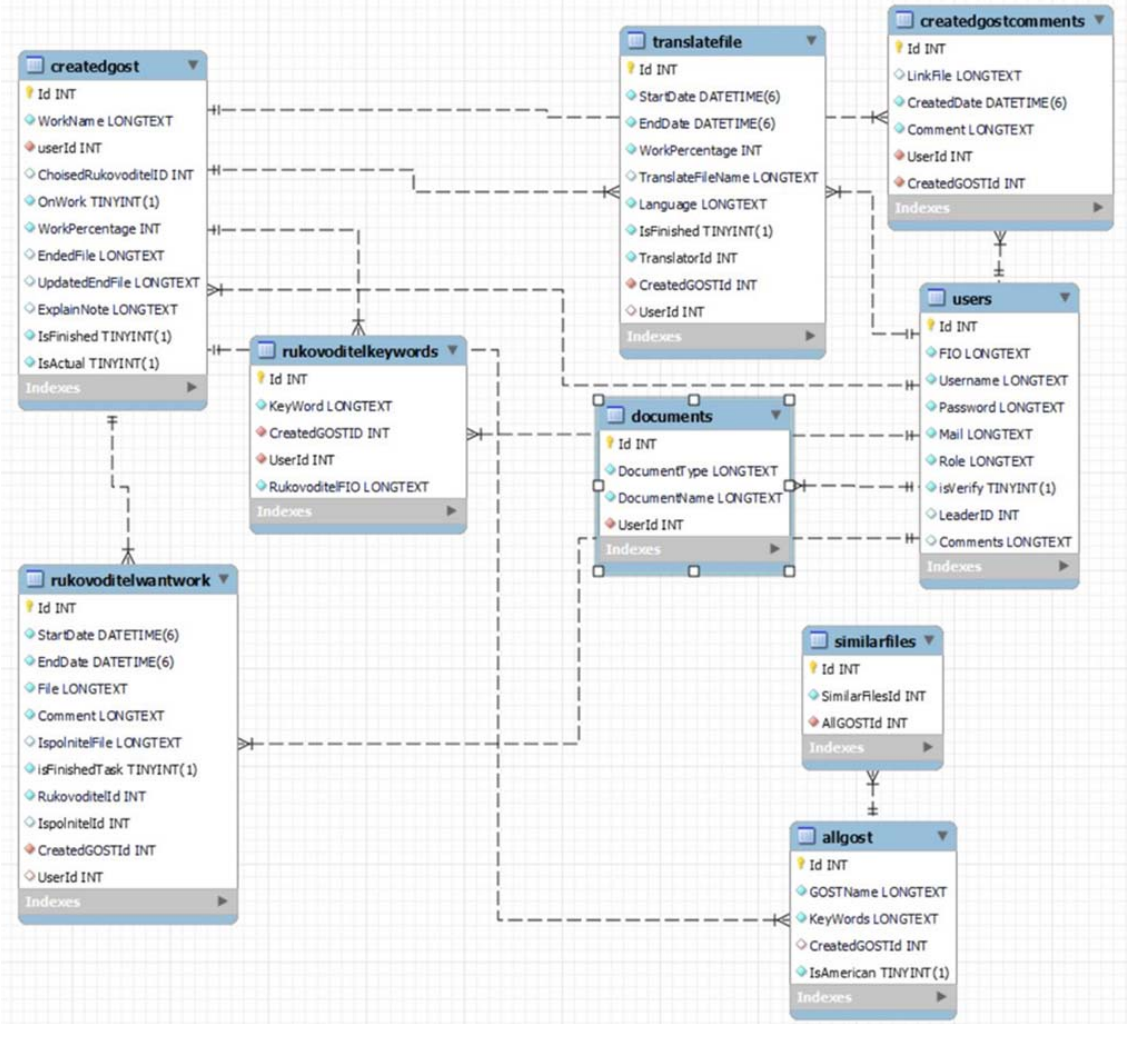

Fig. 2. System Model

Table 1

The important components of the model

Partition	Section Description
Allgost	Keeps a list of files uploaded by both administrators and applicants
CreatedGost	Requests for the creation of standards are stored
CreatedGOSTComments	Contains messages from managers, implementers, translators, administrators, and applicants in the standards creation process
RukovoditelKeyWords	Keywords related to supervisors are stored for selection and recommendation to applicants
RukovoditelWantWork	Presents a schedule in which the manager assigns tasks to performers and leaves work requests
SimilarFile	Files that have the maximum number of similar terms are stored
TranslateFile	Is intended for translators and is provided after the completion of the work by the performers
Documents	It is used to store documents of new users for subsequent verification
User	Contains a list of users of the system

Then, as we work through the use cases, we will gradually refine these responsibilities and eventually arrive at a detailed static model of the system.

5. 2. Harmonization of international standards: classification, extraction, categorization, thesaurus, monitoring, analysis

An important aspect in the harmonization of international standards is the detailed study and correct conversion of units of measurement. The system provides an automatic search in the text of the standard of existing units of measurement and the function of adding new units to the library. In harmonization, three types of standards are defined: identical, modified and non-equivalent. The module for organizing rules for building standards allows to create templates that meet the general requirements for national standards (Fig. 3).

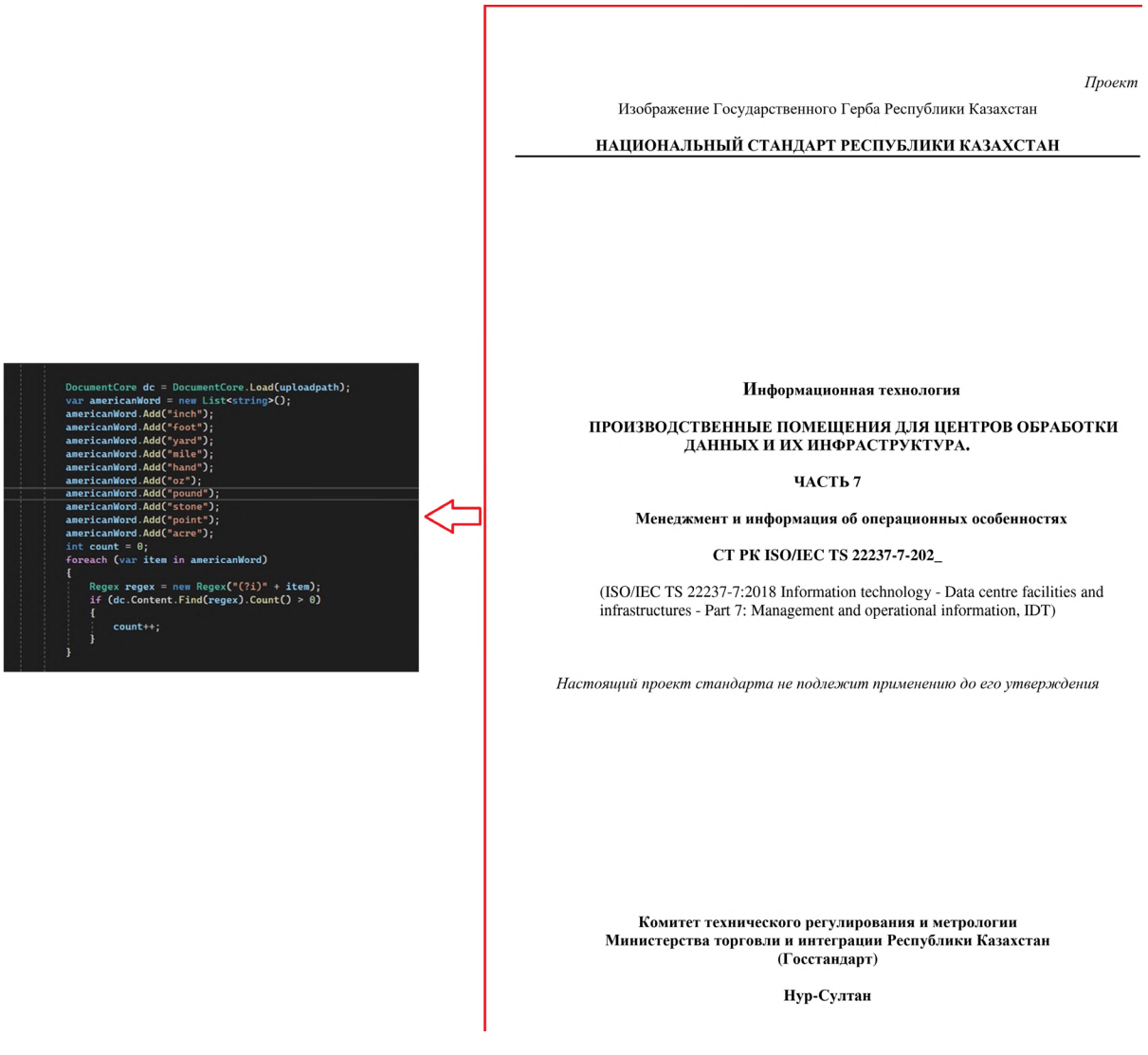

Fig. 3. Automatic formatting Automatic text formatting by template

Examples of the main standards used in the system were extracted from the web resource of the Kazakhstan Institute of Standardization and Metrology [20]. This web resource provides access to draft standards that are under discussion. This made it possible to build a database of standards for integration into the repository (Fig. 4).

After data processing, managers have the opportunity to review the standard for familiarization. At this stage, applicants have the opportunity to get acquainted with current leaders in the relevant field, which is formed on the basis of keywords that determine the direction of activity of each leader. The core module developed in this process is based on quantitative content analysis and the generation of keywords from the source document. These key terms are compared with other documents, and the results of this processing form a database in which distributed key terms from the document text are stored (Fig. 5).

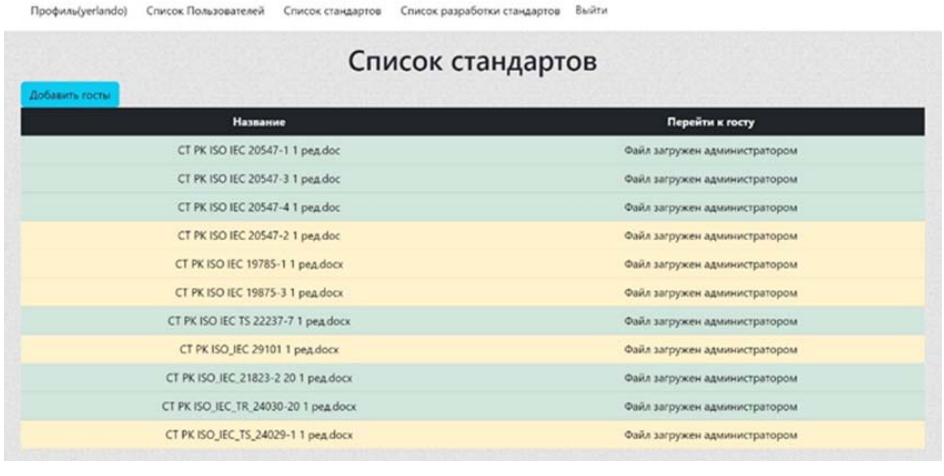

Fig. 4. Search for standards close to keywords

This list is available only to those managers who have expertise in this area. The database of managers also contains keywords that help select experts for various areas.

6. Discussion of the results of the study of mechanisms for monitoring and control of processes in the development of machine-readable standards

In combination with theoretical principles, practical recommendations and proposed methods obtained during the study, a mathematical model the results of which show the main properties of the implementation of the procedure for automatically assigning an updated or harmonized standard to one or more classes. In the same section, Fig. 1 shows a BPMN process diagram, which is divided into three parts: the client, administrator and database interface. The design part was developed taking into account the basic requirements for such tools, where the following results are presented:

– the result of the design part is the developed important components of the model, shown in Table 1. Further, in Fig. 2, the developed class diagram was demonstrated, more precisely, the model of the system on which the database was built, the database structure and functional system were developed, the mechanism for monitoring and managing processes in the development of national and international standards was analyzed;

– concluding the fifth section, Fig. 3–5 show the developed system in the form of the main interface windows, designed in accordance with the main stages of harmonization of international standards.

The application of ontology to standardization provides a number of significant advantages, including the elimination of duplication of terms, the prevention of inconsistencies in terms, the prevention of re-introduction of existing standards, the possibility of scaling standards between organizations and regions, as well as the discovery of hidden knowledge from existing ontologies.

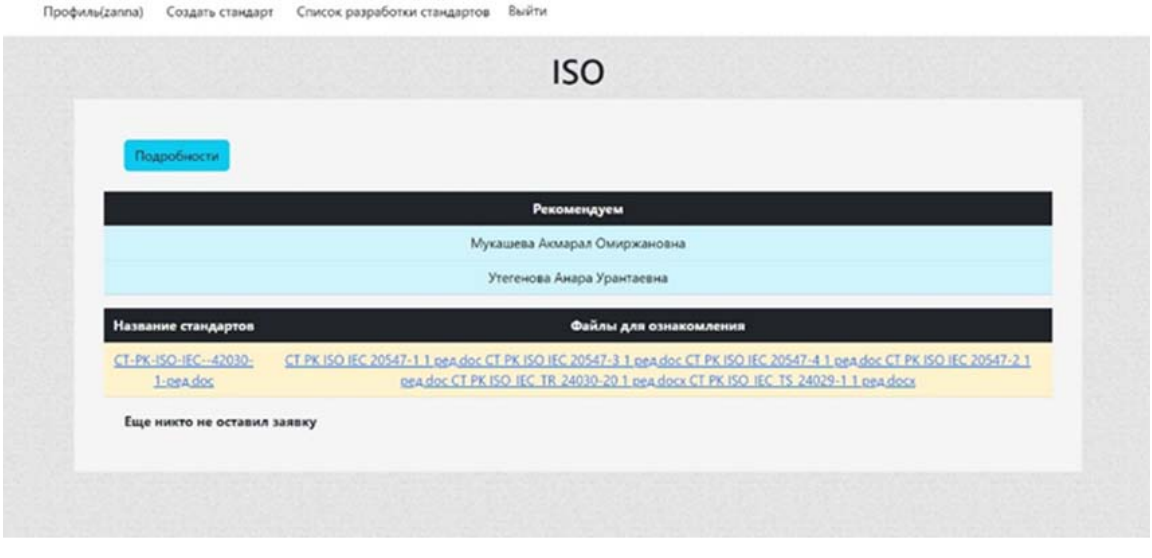

Fig. 5. Search for similar standards using keywords

The obtained stage of harmonization of international standards includes: standard classification model, entity extraction module, categorization module, thesaurus formation module, monitoring and analysis module, developed model of primary transformation of text for ontology formation and subsequent editing by specialists in the profile area, developed model using a systematic approach to methodology ontological design for the formation of an unambiguous formulation of the definitions of harmonized and newly created standards without losing the connection between the entities, developed ontologies according to the standards and analyzed the results of the comparison to identify contradictions or inconsistencies in the terminological unity according to priorities.

Considering the obtained results, it can be concluded that this tool can be used in different industries and business models. The system will provide a barrier-free exchange of information to optimize the use of resources for decision-making at the level of government, public and private sector.

The main drawback of the results obtained is the narrow focus of a carefully researched and developed system of standards specifically for the aerospace and military industries of Kazakhstan. Also, the general lack of clear mechanisms and methods in Kazakhstan for the development and harmonization of national standards by industry makes it difficult to quickly create new or adapted standards.

The application of the developed system in other sectors of Kazakhstan may require additional methods and solutions to create a digital national standard for Kazakhstan. In the future, it is planned to continue research into methods of harmonization and development of national standards for all sectors of Republic of Kazakhstan.

The results obtained are explained by the use of new methods of ontological design, methods of intelligent management in the processes of the operational cycle of standards development and tools of intelligent systems when addressing issues of improving the mechanism for monitoring and managing standards development processes. national standards that pose new scientific and scientific-technical tasks applicable to standardization, as strategically important elements of the country's ecosystem, from an IT point of view, which are based on previously published works of the authors [23–25]. A model has been formed for standard developers using technological integration of optimal methods for creating a single metadata repository in accordance with developed standards and requirements based on expert systems in the form of knowledge base rules, based on previous research by the authors [26].

To date, a sufficient number of knowledge mapping models have been developed. Each of them has its own advantages and disadvantages, and therefore for each specific task it is necessary to choose its own model, continuing the previous studies of the authors [27]. This will determine not so much the effectiveness of the problem as the possibility of solving it as a whole, as demonstrated in the research of the co-authors of this article [28].

The rapid pace of development of new technologies is generating the growth of services and infrastructure of modern standardization, while the problems of lagging behind its performance are increasing. In this regard, it is necessary to reduce the time of work, in some cases it takes 3 years from the beginning of the application to the implementation of the standard, which contradicts the relevance of the results, especially in the technical direction, all the work done also loses its meaning and incurs irretrievable economic losses of public funds [29].

However, after considering these models, it becomes obvious that the creation of human-like intelligence in terms of independent processing of information and use of knowledge is still impossible. At least, not in pure models. It is logical to assume that it is necessary to combine some of these models to obtain the basic qualities characteristic of the human mind. But even in the case of creating such an intelligence, there will still be some qualities that are practically impossible to program – such as extraordinary thinking and the ability to create, the ability to make assumptions and guesses, and create theories and ideas. And, strangely enough, these qualities at the time pushed humanity to the idea of creating artificial intelligence.

The expected results of the study will contribute to the development of the field of "Standardization, Certification and Metrology" (by industry), the development of intelligent systems for monitoring text markup and key data, semantic data analysis, a multi-agent approach to extracting information from texts, etc.

The developed model of requirements for the management of government data in public sector bodies and organizations made it possible to identify and consider the requirements for data management, data description and assessment of the effectiveness of metadata management. Research work was also carried out on the data architecture model, on the example of the Group of National Standards "Information Technologies".

One of the significant aspects investigated in this study was data quality management. It identified ways to identify high-quality data among the total amount of data, which is one of the main objectives of the project.

When managing government data in public sector bodies and organizations, it is necessary to pay attention to the methods of data storage and possible operations with it. In this regard, the basic architecture of the data storage system was proposed, as well as options for its development, taking into account the criteria of data sharing and metadata recovery capabilities.

The operational cycle of the development of the national standard was concretized on the example of the National Standardization Body – RSE "Kazakhstan Institute of Standardization and Certification". Promising areas of the organization's activities, the standardization plan, and the existing technical committees in the territory of the Republic of Kazakhstan, made it possible to identify new areas for research within the framework of research work carried out by the Republic of Kazakhstan Institute of Standardization and Metrology in the field of technical regulation.

At the moment, the entire procedure for harmonizing standards is carried out by e-mail, which is often the only reason for delays in decision-making. The developers send the experts a draft standard in three languages, Kazakh, and English, as well as a cover letter. The enormous workload of examiners is caused by non-compliance with the requirements and the lack of awareness of management about the overall workload of a particular examiner. The uneven distribution of workloads leads to the irrational functioning of the entire system. Speaking about process automation, it is necessary to start from the lower layers of all stages, and only after that to receive real data and make effective, and in some cases cardinal decisions to change the methods of

operation of the existing system. It is necessary to take into account personal responsibility when forming the position of experts and creating a digital footprint of the document to exclude corrupt transactions of interested parties.

Summarizing the history of categorization, it should be noted that the completeness of the use of a particular type depends significantly on the initial requirements for the use of ontologies and the scope of application. In the presented study, the authors came to the conclusion of the development of a formal language of knowledge representations to fully cover the tasks aimed at creating a repository of knowledge on key entities.

According to the generally accepted categorization of ontologies, the following four types were initially distinguished:

1. Ontologies of content for the reuse of knowledge. These ontologies include other subcategories: problem ontologies, domain ontologies, and general ontologies.

2. Ontologies of communication ("say and ask") for knowledge sharing.

3. Indexing ontologies to find cases.

4. Meta-ontologies are equivalent to what other authors call knowledge representation ontologies.

Particular attention is paid to methods of teaching ontology, which reduce the effort in the process of acquiring knowledge; an ontology aggregation that generates a unique ontology from multiple ontologies; Ontology alignment, which establishes different types of mapping between ontologies and an assessment of the ontology's content.

Project documentation has been developed for mechanisms for monitoring and controlling processes in the development of national and international standards, as well as the operational cycle for the development of a standard based on a model using a systematic approach to the methodology of ontological engineering. A study has been carried out and intermediate results have been obtained on the current tasks of analyzing the applicability of the ontological engineering methodology. Project solutions for the implementation of the software agreed with experts in the field of standardization represented by supervising bodies.

Existing solutions in the field of standardization find their application in terms of solving narrower problems, areas that can stimulate the mass introduction of ontotechnologies are highlighted.

7. Conclusions

1. The task of implementing a procedure for automatically assigning an updated or harmonized standard to one or more classes. The solution to the problem is carried out by selecting optimal parameters to obtain the minimum discrepancy between the actions of experts and automatic selection. As a result, a 93 % agreement between the results and the results of manual text processing was achieved. Most documents related to the harmonized document were automatically excluded, and experts were provided with a list of recommended documents for review. This made it possible to reduce the processing time of an average-sized document to 1500 milliseconds. It is important to note that the developed ontologies were designed taking into account reuse criteria;

2. The calculations carried out were developed and implemented as additional software modules to solve the problem of harmonization of international standards, a standard classification model, an entity extraction module, a categorization module, a thesaurus generation module, a module for monitoring and analyzing development stages. An intelligent automated system for managing, monitoring and controlling the processes of the operational cycle of standards currently fully automates data processing. The resulting changes make it possible to monitor in real time the entire standards development cycle and synchronously create a digital footprint.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

This study was carried out and funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (grant No. AP09058441 "Building a system of intelligent control of the processes of development and harmonization of standards within the framework of interstate and national standardization based on ontological engineering").

Data availability

Manuscript has data included as electronic supplementary material.

Acknowledgments

The authors also express special gratitude to the Republican State Enterprise on the right of economic management "Institute of Information and Computing Technologies" and the Almaty University of Energy and Communications named after Gumarbek Daukeev for providing data for scientific research.

References

1. DeNardis, L. (Ed.) (2011). Opening Standards: The Global Politics of Interoperability. The MIT Press, 272. Available at: http://www.jstor.org/stable/j.ctt5hhmcx Last accessed: 07.09.2023
2. Chirumalla, K. (2021). Building digitally-enabled process innovation in the process industries: A dynamic capabilities approach. Technovation, 105, 102256. doi: https://doi.org/10.1016/j.technovation.2021.102256
3. Czarny, D., Olsen, H. et al. (2020). Project 2 Standards of the Future – Pilot Petroleum sector. Internal CCMC-Report.
4. Czarny, D., Salcedo, I. et al. (2020). Classification Sheme and Use Cases. Internal IEC SG 12 Report.

5. Czarny, D., Wischhöfer, C. (2020), Aktuelles zum Thema SMART Standards – Die IDiS-Initiative. DIN-Mitteilungen. Berlin: Beuth-Verlag, 5–9.
6. Esser, M., Koch, H., Willuhn, H. (2017). Digitale Content-Dienstleistungen aus dem zentralen XML Content Repository – Zentrale Ablage von Inhalten und Trennung der Inhalte von ihrer Darstellungsform. DIN-Mitteilungen. Berlin: Beuth-Verlag, 18–23.
7. Huang, T. (2020). Design and Key Technology Research of Civil-Military Integration Regulations and Standards Management System. Smart Innovation, Systems and Technologies, 201–207. doi: https://doi.org/10.1007/978-981-15-5073-7_21
8. Ryu, H.-E. (2021). Standards Harmonization and Asymmetric Compliance Technology. Journal of Korea Trade, 25 (3), 1–20. doi: https://doi.org/10.35611/jkt.2021.25.3.1
9. Harmonised Standards. Available at: https://single-market-economy.ec.europa.eu/single-market/european-standards/harmonised-standards_en
10. Chaisricharoen, R., Srimaharaj, W., Chaising, S., Pamanee, K. (2022). Classification Approach for Industry Standards Categorization. 2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), 308–313. doi: https://doi.org/10.1109/ectidamtncon53731.2022.9720348
11. Standards in Europe (2023). Available at: https://europa.eu/youreurope/business/product-requirements/standards/standards-in-europe/index_en.htm Last accessed: 01.10.2023
12. Ford-Baxter, T., Faulkner, K., Masunaga, J. (2022). Situating Information Literacy: A Case Study Exploring Faculty Knowledge of National Disciplinary Standards and Local Program Learning Outcomes. The Journal of Academic Librarianship, 48 (3), 102523. doi: https://doi.org/10.1016/j.acalib.2022.102523
13. Barthelemy, F., Cochez, M., Dimitriadis, I., Karim, N., Loutas, N., Magnisalis, I. et al. (2022). Towards a standard-based open data ecosystem: analysis of DCAT-AP use at national and European level. Electronic Government, an International Journal, 18 (2), 137–180. doi: https://doi.org/10.1504/eg.2022.121856
14. Reategui, E., Bigolin, M., Carniato, M., dos Santos, R. A. (2022). Evaluating the Performance of SOBEK Text Mining Keyword Extraction Algorithm. Machine Learning and Knowledge Extraction. Berlin, Heidelberg: Springer-Verlag, 233–243. doi: https://doi.org/10.1007/978-3-031-14463-9_15
15. Mehrabi, E., Mohebi, A., Ahmadi, A. (2021). Improved Keyword Extraction for Persian Academic Texts Using RAKE Algorithm; Case Study: Persian Theses and Dissertations. Iranian Journal of Information Processing and Management, 37 (1), 197–228. doi: https://doi.org/10.52547/jipm.37.1.197
16. Miah, M. S. U., Sulaiman, J., Azad, S., Zamli, K. Z., Jose, R. (2021). Comparison of document similarity algorithms in extracting document keywords from an academic paper. 2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM). Pekan, 631–636. doi: https://doi.org/10.1109/icsecs52883.2021.00121
17. Yang, Y., Wu, Z., Yang, Y., Lian, S., Guo, F., Wang, Z. (2022). A Survey of Information Extraction Based on Deep Learning. Applied Sciences, 12 (19), 9691. doi: https://doi.org/10.3390/app12199691
18. Nie, B., Sun, S. (2017). Using Text Mining Techniques to Identify Research Trends: A Case Study of Design Research. Applied Sciences, 7 (4), 401. doi: https://doi.org/10.3390/app7040401
19. Schacht, M., Wischhöfer, C. (2014). Nutzen semantischer Technologien in der Normung und Anwendung - Das Heben eines Wissensschatzes: Normen semantisch analysieren. Inhalte zielgerichtet extrahieren und in Folgeprozessen verwenden. DIN-Mitteilungen Oktober 2014. Berlin: Beuth Verlag, 6–11.
20. Discussion of ST RK projects. Available at: https://www.ksm.kz/public-discussion/standards/discussion-of-draft-of-st-rk/ Last accessed: 01.10.2023
21. Baibolov, A., Sydykov, S., Alibek, N., Tokmoldayev, A., Turdybek, B., Jurado, F., Kassym, R. (2022). Map of zoning of the territory of Kazakhstan by the average temperature of the heating period in order to select a heat pump system of heat supply: A case study. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44 (3), 7303–7315. doi: https://doi.org/10.1080/15567036.2022.2108168
22. Kassym, R. T., Taldybayeva, A. S., Omar, D. R., Alibek, N. B., Kuder, K. M., Isakhanov, M. Z., Omarov, R. A. (2021). Experimental results of functional characteristics of IOT for free range sheep breeding. International Journal of Agricultural Resources, Governance and Ecology, 17 (2/4), 301–315. doi: https://doi.org/10.1504/ijarge.2021.10044149
23. Utegenova, A., Naumenko, V., Alenova, R. (2021). Author's certificate No. 610144604. Ontological Engineering-Based Standards Development Process Management System.
24. Baimuratov, O. A. (2022). Copyright certificate No. 27104. WEB Scraping of ISO standards. 13.06.2022.
25. Baimuratov O. A. (2022). Author's certificate No. 27082. Collector of standard data of Kazakhstan. 10.06.2022.

26. Bapyshev, A. M., Tulendi, A. A. (2022). Harmonization of Standards Taking into Account the Automatic Verification of Correctness and Consistency of Terminology Based on Machine Learning Technologies. Development of Education and Science in New Kazakhstan. Nur-Sultan, 189–192.
27. Utegenova, A. U., Suimenbayeva, Zh. B., Bapyshev, A. M., Áden, A. E., Sakan, E. M. (2022). Methods for extracting terms from normative documentation based on ontology. Problems of optimization of complex systems. XVIII International. Almaty, 174–181.
28. Utegenova, A., Ermoldina, G., Bapyshev, A., Naumenko, V., Aden, A. (2022). Ontological model of knowledge representation and organization, taking into account the recommendations of international standards in the educational process of the military university. KazATC Bulletin, 123 (4), 307–318. doi: https://doi.org/10.52167/1609-1817-2022-123-4-307-318
29. Utegenova, A., Ermoldina, G., Bapyshev, A., Naumenko, V. (2023). System of intelligent control of processes of development and harmonization of standards in the field of military and space activities on the basis of ontological engineering. KazATC Bulletin, 127 (4), 148–159. Available at: https://vestnik.alt.edu.kz/index.php/journal/article/view/1315
30. Kassym, R., Baymukhanbetov, A., Ermolina, G., Isakhanov, M., Tokmoldaev, A. (2023). Environmental aspects and impact on environmental changes of some useful elements and hazardous heavy metals in sheep and cow manure. KazATC Bulletin, 126 (3), 479–489. doi: https://doi.org/10.52167/1609-1817-2023-126-3-479-489