Modernizing quality management with formal languages and neural networks

Irbulat Utepbergenov, Shara Toibayeva
Department of Automation and Control, Gumarbek.Daukeev Almaty University of Power Engineering and Telecommunications, Almaty, Kazahstan

Article Info

Article history:

Received Aug 13, 2024
Revised Apr 1, 2025
Accepted May 24, 2025

Keywords:

Automated control system
Document structure
Formal language
Intelligent system
Quality management

ABSTRACT

This paper explores the integration of formal languages and neural networks into quality management systems to enhance efficiency and sustainability. Formal languages standardize regulatory documents, reducing misinterpretation and simplifying modification, contributing to innovative infrastructure (SDG 9). Recurrent neural networks (RNNs) automate document analysis, non-conformance detection, and decision-making, improving production efficiency and promoting responsible consumption (SDG 12). Automation in quality management reduces costs, enhances competitiveness, and aligns with decent work and economic growth (SDG 8). Standardizing documentation and automating quality control enhance workforce competencies and support quality education (SDG 4). These technologies strengthen regulatory transparency, reduce legal risks, and improve governance, supporting strong institutions (SDG 16). The proposed approach fosters sustainable development through digitalization and automation, ensuring efficiency, innovation, and compliance with environmental and social standards.

Corresponding Author:

Shara Toybayeva
Department of Automation and Control, Gumarbek.Daukeev Almaty University of Power Engineering and Telecommunications
Baytursynuli 126/1, Almaty, Kazahstan
Email: sh.toibaeva@aues.kz

1. INTRODUCTION

The economic potential of Kazakhstan is directly related to the development of business, and the management environment affects the efficiency and effectiveness of the enterprise. By joining the WTO, Kazakhstan has increased competition between enterprises to occupy its niche in the market and strengthen its position, which requires the release of higher quality products to the market, capable of satisfying the end users [1]. The enterprise adapts to changes in the external environment, which determines its organizational component of management, which includes many links (units), classified by appropriate strategies, relationships (relationships), and documented requirements.

Today, the effective result of the organization of all functions of the enterprise: planning, organization, production, control, work with personnel, etc. is characterized by the concept of quality. For example, the quality of a product reflects the quality of its design, development, manufacturing, and sales processes. Thus, quality management affects the entire product life cycle: from research in marketing, research and development, design stages, and up to the production of products, sales and maintenance. In this regard, the current steps in the development of quality improvement techniques cover both the problems of improving the quality of products or services, and the quality of management itself [2].

The main task for effective management are the stages of formation, implementation and use of an automated system of quality management. The control system should provide access to the documentation of the enterprise, quality output of the requested information to take management actions to solve the problems of the enterprise relevant at a certain moment. At product quality management, the objects of management are processes, and managing decisions are formed by comparing the indicators of the actual state of the managed process with its planned ones. To successfully perform its functions, an enterprise needs to manage significant correlated activities as a system of processes. According to the results of the analysis of literary sources for the last 20 years, we can say that in the world, in particular in Kazakhstan, the scientific direction related to the problems of implementation and automation of quality management systems, technology analysis of business processes of various organizations is developing.

Turning to statistics, we can note that the starting point of management development in Kazakhstan are the 90's, when the management situation and the market system have changed dramatically, and there was a need to study the foreign experience of management, to apply it in the country. More and more interest in the implementation of the QMS and automation of business processes of enterprises is noted. For example, the query “automation of quality management” in the information base of scientific electronic library eLIBRARY.RU received 43085 results [3], [4], which indicates the interest in quality management systems and its automation in the scientific community.

In an industrialized country, it is necessary to meet the challenges of implementing and keeping up-to-date modern management systems, where competition, knowledge-intensive, innovative and technologically sophisticated production is developed. Since the beginning of 2002, the organization, design and implementation of QMS of organizations in Kazakhstan has been noticeably strengthened [5], [6]. In the field of quality management, the orders of Prime Minister of the Republic of Kazakhstan No 28-r dated 06.02.2004 and No 175-r dated 27.06.2006 on the soonest rearrangement of enterprises according to ISO standards have been adopted, so on achieving the set goals in the country the proper material and technical support, normative and methodological base for implementation of international standards is created [7], [8]. The regulatory framework of the Republic of Kazakhstan [7], [9] includes 36 state standards, where international ISO standards were taken as a basis and adopted as state standards of the Republic of Kazakhstan. Development and implementation of standards is part of the standardization plan of the Republic of Kazakhstan each year in the field of management.

In their work, Garrido-Moreno *et al.* [10] argue that the companies are forced to react quickly and rethink their business strategies in an uncertain and complex economic environment. Innovation has become a strategic imperative for adapting to market changes and maintaining competitiveness, and resilience is gaining attention as a prerequisite for organizations to successfully respond to external environmental pressures. Innovative companies are adopting new systems management approaches, and enterprise quality managers are achieving greater results for their organizations in less time [11], [12].

Several problems can be identified that quality management professionals encounter when using such static management systems:

a. Unproductive use of work time. Managers have to enter the same data several times to collect data from different sources and combine it into single reports.
b. Lack of real-time, relevant information. It takes some time to distribute and update information throughout the organization. The data updates slowly in the system, and in order to provide an accurate assessment of the situation, the most recent changes in compliance data must be processed.
c. Limiting access to accurate analysis. Facility management relies on managers to perform surveillance of what is going on in the facility. Since the managers themselves are not quality managers-they rely on the latter to analyze and present data they can understand and read. It takes a lot of time to create reports, thereby slowing down the feedback process.
d. Preparing monthly reports for most businesses is a duplication of data. When calculating data, it takes time for managers to submit reports to supervisors.
e. Ineffective communication with employees. Creating connections with employees and training people is very time consuming. Quality managers spend a lot of time on compliance issues without focusing on other activities.
f. Lack of databases of documented information and as a result, no automated decision support system. Lack of prompt notification and response to the facts of deviations of the values of the planned indicators.

Consequently, the problem lies in the lack of efficiency of traditional quality management systems due to the complexity of regulatory documents, subjectivity of their interpretation and dependence on the human factor. The lack of automated tools for processing, analyzing and updating regulatory documents leads to errors, delays and increased costs in production processes. To solve this problem, we propose the integration of formal languages and neural network models, which provide accurate structuring of regulatory data, automate their processing and increase the objectivity of decision-making in the quality management system.

2. THE IMPORTANCE OF BUILDING A FORMAL LANGUAGE

A formal language is treated as a set with all strings under a finite number of alphabets. The formal language theory states that the study of all families is done by formal methods. This formal method is mainly used to develop and verify software and hardware systems. It is usually described for formal languages and automata theory. Typically, formal languages deal with a set of strings and are defined in two sets of rules. The first rule deals with syntax, which specifies how to use symbols, while the second rule deals with semantics, which tells the meaning of symbols and legal expressions [13].

Constructing a formal language that characterizes normative documents is an important tool for ensuring accuracy and clarity in describing the rules and requirements contained in normative documents. Regulatory documents are used in various fields of activity, from jurisprudence to technical documentation, and often have legal significance. Therefore, they need to be written in clear and comprehensible language so that users can easily understand the requirements they contain. However, normative documents are often complex and confusing, which makes them difficult to understand and leads to errors in their interpretation and application. Building a formal language allows you to develop a common language to be used to describe documents and establish clear syntactic rules for its use. This will make it easier for the user to understand the requirements contained in normative documents and reduce the likelihood of errors.

In addition, building a formal language can help in developing new regulatory documents and amending existing ones. With the language, you can quickly identify what concepts are already used in other normative documents, what syntax rules apply, and how different requirements are described. This can greatly speed up the process of developing and amending documents [14]. Thus, in order to increase the speed of creating documents and the accuracy of the data contained in them, automated document management systems are being actively developed and widely used. The development of a universal formal language will make it possible to create a system of automated formation of documents of any composition and purpose.

3. BUILDING A FORMAL LANGUAGE

The construction of a formal language characterizing normative documents is an important tool for ensuring accuracy and clarity in the description of rules and requirements contained in normative documents. Normative documents are used in various fields of activity, from jurisprudence to technical documentation, and often have legal significance. Therefore, it is necessary that they be written in a clear and understandable language so that users can easily understand the requirements contained in them.

However, normative documents are often complex and confusing, which makes them difficult to understand and leads to errors in their interpretation and application. The construction of a formal language makes it possible to develop a single language that will be used to describe documents, and to establish clear syntactic rules for its use. This will make it easier for the user to understand the requirements contained in normative documents and reduce the likelihood of errors.

A formal language is considered as a set with all strings which are under a finite number of alphabet. Formal language theory states that studies of all families on formal methods. This formal method is mainly used for developing and verifying the software and hardware systems. It is typically described for formal languages and automata theory. Generally formal languages deal with the set of strings and defined in two set of rules. The first rule is about syntax which tells that how to use the symbols whereas the second rule is about semantics which tells that meaning of the symbols and legal expressions [15], [16].

The system based on the formal language is fundamentally different from its counterparts in that it implements a new approach in terms of documenting, which consists in constructing the document plan from special components that are computer models of the document's constructive assemblies. This process is called “document construction” [17]. Using FL, you can describe the structure of almost any document. The structure of the order in the above example in Figure 1 can be described by a set of the following nonterminal symbols of the F

Building a formal language can help in the development of new normative documents and amendments to existing ones. The language can be used to quickly determine what concepts are already used in other normative documents, what syntax rules apply, and how various requirements are described. This can greatly speed up the process of developing and amending documents [18].

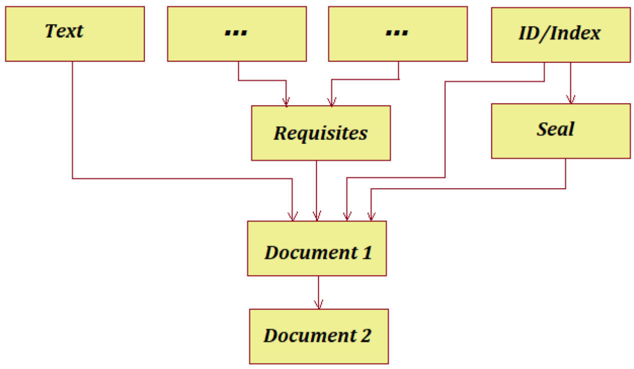


Figure 1. An example of using FL to describe the structure of a document

The nonterminal symbol <Document 1> denotes the entire document. It is part (as a fragment) of a more complex document, indicated in the figure by the symbol <Document 2>;

The <Text> symbol denotes a block of text that can include a set of all characters from the ASCII (or Unicode) table;

The symbol <Requisites> expresses a requisites of a document(name of the position of the person who signed the document, signature and transcript of the signature).

The three-dot (...) denotes other characters of the FL: <Text>, <East Lit>, <<Word>, <Digit>, <Punctuation Mark>, <Ending Mark>, <Letter>, <Arithmetic>.

The normative document can be presented in CT in three forms:

$$CD = (K_{Doc}, B_{Doc}, N_{Doc}), \quad (1)$$

where K_{Doc} is a set of components reflecting the structural elements of the documents; B_{Doc} is a set of links between the structural elements of the documents; N_{Doc} is a set of nodes located at the connection points of the structural elements of the documents.

In turn, the K_{Doc} component set contains three subsets:

$$K_{Doc} = \{K_{DI}, B_{DP}, N_{DZ}\} \quad (2)$$

where K_{DI} - set of information converter components; K_{DI} - set of information converter components; K_{DZ} - final component (“document”).

K_{Doc} topology and coordinate system components are not supported by S_1 boiler. S_1 bylaws have the following types: span; span; logic; zhol; array; and derek and file [19]–[21].

Components are the source of information about the K_{DI} set as shown in Figure 2 to set the source information and transfer it to other components or directly to the document. The algorithmic model and process of the information source component are characterized by the following sequence:

a. Component parameter values and source information to be included in the document;
b. The component stores the specified parameter values and source information in special variables;
c. The component is linked to other components (by means of information contact n1);
d. The document creation process begins;
e. The component transmits the information previously stored in the variable to info contact n1 via info channel a1;
f. The components connected to the component via infocontact n1 receive the information, then process it and send it to the document or send it to the document immediately [22].

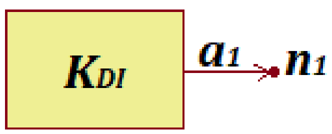


Figure 2. Graphical representation of the source component

Components are transducers of information in the K_{DP} set as shown in Figure 3 for receiving source information, transforming it and transmitting it to other components or directly to the document. Components of this type may have one or more inputs and only one output info-contact. The end component “Document” of the K_{DZ} set as shown in Figure 4 is used to The end component “Document” of the set K_{DZ} as shown in Figure 3 is used to collect all the information coming to it and to form a document.

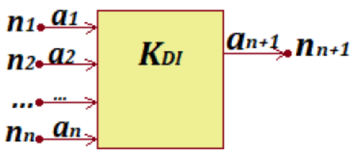


Figure 3. Graphic representation of the information-converter component

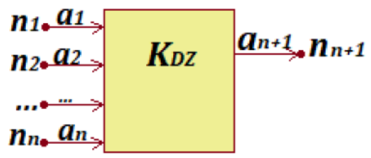


Figure 4. Graphic representation of the “document” component

It can also serve to transfer the formed fragment to another document. Using this component makes it possible to form complex documents consisting of a number of documents (or their fragments). The “Document” component can have one or more input info-contacts (n1...nn) and one output info-contact (nn+1). Thus, the use of formal language is one of the key tools for digitizing the regulatory environment and optimizing management processes [23], [24].

4. MODEL OF NEURAL NETWORK FOR CONVERSION OF WORD INTO MATRIX VECTORS

The bag-of-words (BoW) method represents documents as vectors, where each element corresponds to a unique word in the dictionary and the value of the element indicates the frequency of that word in the document. However, this method does not take word order into account, which can lead to loss of context, as it simply creates a “bag” of words with no relationship between them. To improve context consideration, the N-gram model is used, which extends BoW by considering combinations of N consecutive words. N-grams allow the model to take into account the probability of a word's occurrence based on previous words, making the representation of text more context sensitive. This creates a more complex but also more informative representation of the data, improving the machine's perception of language. The BoW model can be mathematically represented as a matrix, where the rows correspond to documents and the columns to unique words. An example of building such a matrix using BoW is usually illustrated in a diagram (e.g., Figure 5) that visualizes the distribution of words across documents.

	about	bird	heard	is	the	word	you
About the bird, the bird, bird bird bird	1	5	0	0	2	0	0
You heard about the bird	1	1	1	0	1	0	1
The bird is the word	0	1	0	1	2	1	0

Figure 5. Assembly of the bag of words matrix (bag of words)

Computation-based models such as BoW and N-grams, although simple and efficient, are unable to adequately account for the semantic and syntactic aspects of language. This limitation encourages researchers to develop more sophisticated and context-dependent methods for converting text into matrices. The current approach to natural language processing (NLP) [25] makes extensive use of vector-based word representations that place words in a fixed and reduced multidimensional space that reflects the semantic relations between them. Such methods include Word2Vec and global vectors for word representation (GloVe).

Word2Vec includes skip-gram and continuous bag of words (CBOW) models, which represent words as continuous vectors, where the distance between vectors corresponds to the semantic similarity of words. The Skip-Gram model predicts contextual words based on a given central word, whereas CBOW works the other way around by predicting the central word from the context. These models are trained using a two-layer neural network that optimizes word embeddings in such a way as to minimize the logarithm of the error probability in predicting the context or central word. Figure 6 may show an architecture or loss minimization process that reflects how the models are "stretched" to fit the probability distribution of the context words given the central word.

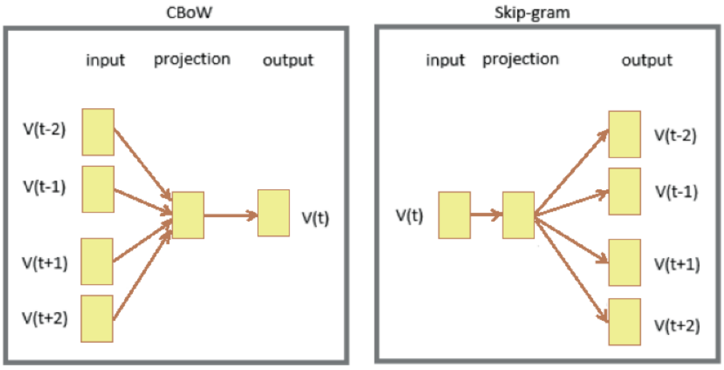


Figure 6. Method for converting words to matrix vectors (Word2Vec)

The advanced age of deep learning has greatly enhanced natural language processing capabilities through innovative models such as bidirectional transformers (BERT) and pre-trained generative transformers (GPT). These state-of-the-art models utilize a transformer architecture that allows the creation of dynamic vector representations of words or phrases, given the context of words, both before and after the target word in a sentence. As a result, the models are able to understand and account for the semantic and syntactic features of the language, which allows for more accurate prediction of word sequences in text. For example, BERT utilizes bidirectional learning by processing text from both left to right and right to left, which helps to capture context in both directions. GPT, on the other hand, focuses on predicting the next word in a sentence using the previous context, which helps generate coherent and grammatically correct text. These models significantly improve language comprehension and processing by taking into account its contextual dependencies and offering more accurate and useful predictions.

$$(\Delta W + h_{max})^n = -\alpha \sum_{k=0}^{n} t \frac{\partial L}{\partial h} \cdot \frac{\partial h_n}{\partial W_k} \tag{3}$$

where, W-gradient update for the recurrent weight; h-the learning rate; L-cost function; t-length of the input circuit; α-regulation coefficient.

In our analytical approach in digital linguistics and computational intelligence, an important component is the transformation of raw textual data into a format that can be efficiently processed using recurrent neural networks (RNNs). This transformation process is necessary because of the complex and dynamic nature of human language, which includes a wealth of sequential and temporal dependencies. RNNs serve as a key tool for analyzing these dependencies due to their ability to remember and incorporate previous information while processing current data. The memory and context preservation mechanisms present in RNNs allow models to capture and analyze sequences of words and phrases in text, making them ideal for natural language processing (NLP) tasks. As a result, machine learning and NLP bridge the humanities field of language learning with modern digital technologies, enabling complex tasks such as machine translation, speech recognition, and automatic text summarization.

The systems enterprise begins with a basic step-the tokenization process. This algorithmic mechanism slices raw text into a series of discrete symbols-words, phrases, or characters-that act as atomic units, forming the building blocks of the upcoming digital transformation. Subsequently, these tokens are further refined using normalization techniques, including lemmatization to separate words into their underlying entities to provide a homogeneous input landscape that mitigates noise and redundancy, thereby preparing the text data for digital coding.

This theoretical framework represents a novel approach that goes beyond simple single-unit encoding and projects tokens into a multidimensional vector space. Rather than simply converting text into numbers, this method creates a multidimensional representation that captures complex patterns of semantics and syntax that are inaccessible at the level of individual linguistic units. By using methods such as Word2Vec, GloVe, and FastText, and training models on large amounts of text data, it is possible to create vector spaces that accurately capture the semantic relationships and syntactic nuances of words. As a result, words with similar meanings are placed close together in this space, creating a more detailed representation of the language and its features.

5. IMPLEMENTATION OF THE PROCESSING PROGRAM

Diving into the proposed sophisticated software development module, the program is carefully built around the use of complex document processing system and the potential of recurrent neural networks, being the core algorithm to drive a conversational intelligent system. Codebase, as a symbol of advanced library and information science, combines interactive querying capabilities on uploaded documents using a sophisticated neural network architecture to analyze documents and create a clear, sensitive processing system for semantic information retrieval. The recurrent neural network, embedded at the heart of AI advances, is becoming the underlying computational architecture for creating interactive experiences from static textual data encapsulated in uploaded documents. The program, based on a sophisticated software development module, has been carefully designed as an interface between the human user and a vast ocean of guided information to persuade him to uncover from the depths of comprehensive document data. By adopting the skill of RNN, the system aims to promote an intelligent, comprehensible communication mechanism capable of decoding and understanding the semantics twisted in the textual corners of each document, and retrieving accurate information in response to user queries.The recurrent neural network algorithm used in the program is shown Figure 7.

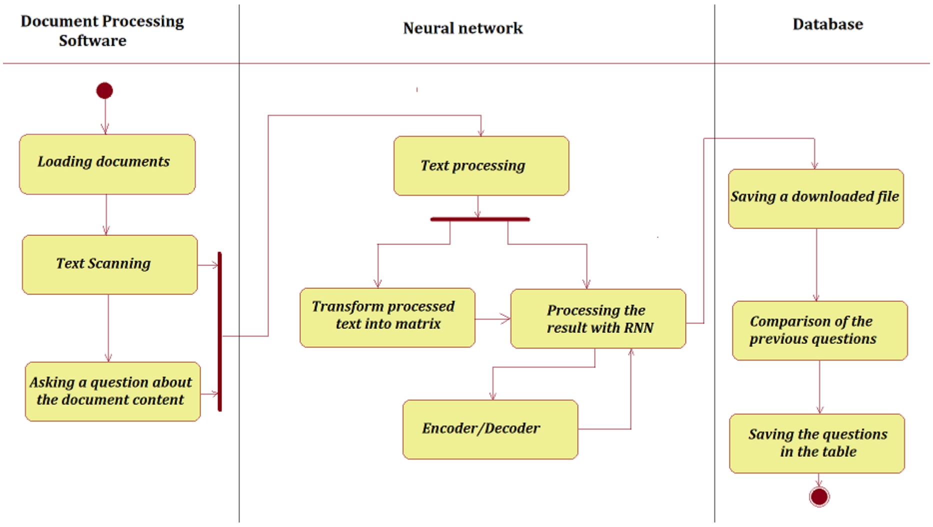


Figure 7. Algorithm for processing a file through a text input neural network

The recurrent neural network is an evolutionary step in the field of deep learning algorithms uniquely adapted to recognize and decipher patterns in sequential data, making it unsurpassed for processing and predicting text sequences.The architecture of a recurrent neural network is characterized by its cyclic neural connections that allow information to flow in a loop, allowing the model to maintain an internal state that reflects the effects of previously encountered data points. This feature gives the recurrent neural network the ability to store historical information and establish important relationships with the sequence, making it more important for tasks such as natural language analysis and time series analysis.

In action, our recurrent neural network software goes through a series of complex steps, gradually Converting user-supplied documents into a coherent, query-responsive system. The pipeline begins with the processing stage, where the function" get_doc_text" ensures that raw text is removed from the uploaded text files. Subsequently, segmentation of text into manageable text chunks is facilitated by the "get_text_chunks" function, which acts as a prior to develop understandable blocks of data ready for neural consumption. After authorization, the user gets to the main page of the program, can browse the software on the main page as shown in Figure 8.

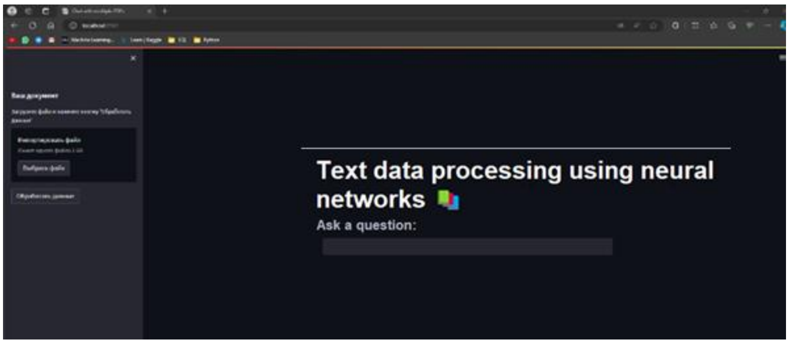


Figure 8. Main page of the program

Turning to embedding, the core of our system's informational understanding lies in the sophisticated text embedding techniques adopted to give meaning to text segments. In essence, with library and informatics, the device integrates the hugging Face library to generate training embeddings using the “hkunlp/instructor-xl” model, evidence of an emerging machine learning interface that extends the semantic understanding of the neural network. These embeddings, which are neural representations of text, close the gap between central sublingual structures and a structured quantitative understanding of the processing capabilities of a recurrent neural network. In the left window we can import the desired file, for this purpose we select the desired file in the window that appears by clicking on the button” Choose file”. In the window that opens, select the desired file and import it. Having selected the file we need, click the “Open” button and wait for it to load. After loading, it is displayed that it is loaded in our program. Once again, we check and press the button “Create data” as shown in Figure 9. Then the processing of the result begins.

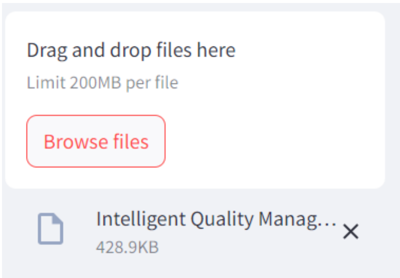


Figure 9. Editing imported files

In our quest to improve the dialog system, the program achieves significant success by creating a context-aware and seamless network of interactions. A conversation retrieval chain based on the Large Language Model and utilizing the HuggingFaceHub model, including the “Flan-T5 Extensible Large”, equipped with a powerful memory module for storing dialog history, allows context to be stored in a conversation. This scheme helps the model extract relevant information from embedded texts, ensuring accurate responses to user queries and maintaining a natural flow of dialog.

With user-centricity in mind, the program applies the Streamit library to create an intuitive graphical user interface (GUI). This interface allows users to upload PDF documents, edit them, and interact through a chat format to get the information they need. With a robust HTML model, the interaction between users and bots becomes visually clear, simplifying complex computational processes and providing access to information for inexperienced users.

The system, along with advanced recurrent neural networks, ushers in a new era of document analysis and interaction by offering users the ability to easily query and discuss document content. By combining sophisticated machine learning algorithms with intuitive human interaction patterns, the recurrent neural network in the Library and Information Science system not only decodes the semantic power of data, but also brings statistical documentation to life, transforming it into a dynamic and intelligent dialog module at the forefront of AI advances. In the context of uploading a document as a data source and then validating the responses generated by the system, several systematic methodologies can be used, offering a broad period to validate the performance of the system against several criteria.

Given the different character sets and encoding format, confirming that the loaded data correctly interprets these elements ensures seamless processing, especially for non-ASCII characters or complex print scenes. Testing should include verifying that special characters are present and correctly encoded to avoid misinterpretation or loss of information during the natural language processing pipeline. The software detected all errors and proceeded to correct them as shown in Figure 10.

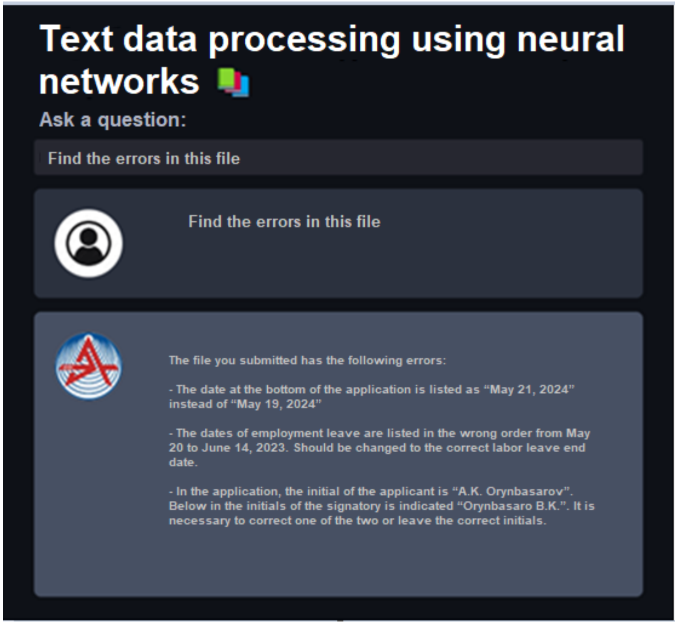
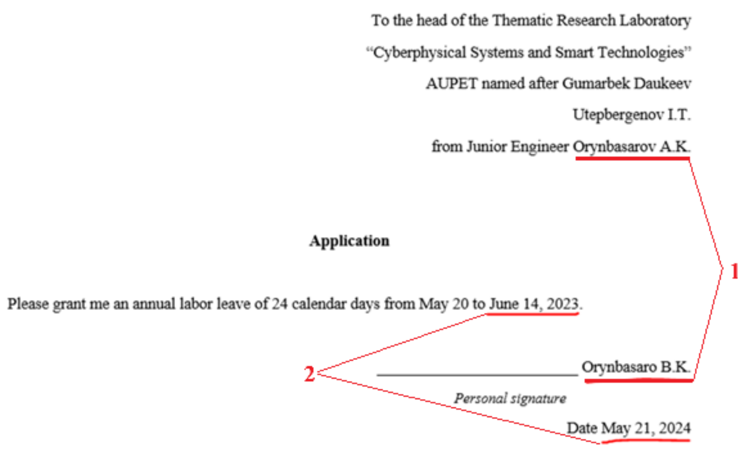


Figure 10. Finding inconsistencies in the document

Once the data has been loaded, tokenization marks the next critical point for validation. This process, which is crucial when preparing text for the NLP pipeline, must be carefully checked for accuracy, ensuring that there are no tokenization anomalies such as unexpected character separation or incorrect handling of sentence boundaries. The methodological approach involves comparing the tokens generated by the system to human-generated controls or to predefined tokenization benchmarks. This verification process ensures that the punctuation, distribution, and token generation of the NLP system meet the requirements of linguistic standards and tasks as shown in Figure 11.


To the head of the Thematic Research Laboratory
"Cyberphysical Systems and Smart Technologies"
AUPET named after Gumarbek Daukeev
Utepbergenov I.T.
from Junior Engineer Orynbasarov A.K.

Application

Please grant me an annual labor leave of 24 calendar days from May 20 to June 14, 2023.

_______________ Orynbasaro B.K.
Personal signature
Date May 21, 2024

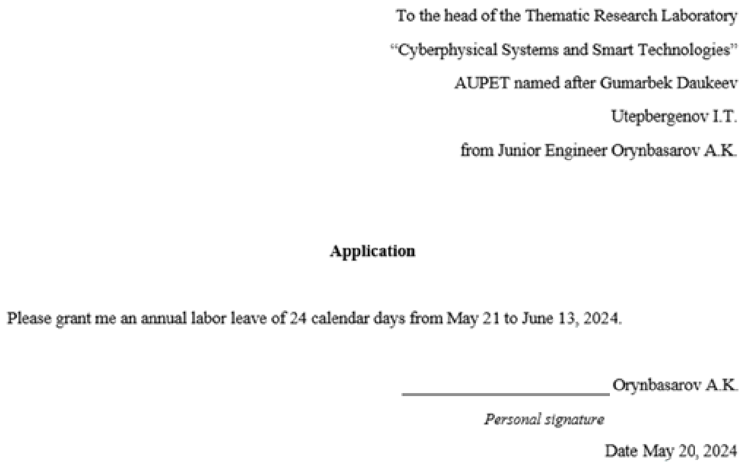
Figure 11. Marking a discrepancy in processing

The document contains the following errors:

a. The beginnings of the author of the document are not correct;
b. The dates of the application are incorrect. The body of the application states a leave of absence from May 20, 2024 to June 14, 2023 (i.e., the end date of the leave of absence is before the start date). In addition, at the end of the statement, the delivery date is proposed as May 21, 2024, which also does not agree with the end date of the leave in 2023).

After processing all the listed errors we will get a fully processed document at the output. When comparing it is worth checking the corrections as shown in Figure 12. We can see the changes marked accordingly. Each correction has been made with logical and chronological consistency in mind, which helps avoid any contradictions in the final version.

The integration of artificial neural networks opens up new horizons in adaptability and improves the efficiency of production systems. Due to their ability to learn, analyze large amounts of data, search for hidden dependencies, and perform predictive calculations in real time, they ensure the most accurate and timely management decisions. Thus, artificial neural network technologies form the foundation for building intelligent, self-adapting, and highly reliable control systems in modern industry.


To the head of the Thematic Research Laboratory
"Cyberphysical Systems and Smart Technologies"
AUPET named after Gumarbek Daukeev
Utepbergenov I.T.
from Junior Engineer Orynbasarov A.K.

Application

Please grant me an annual labor leave of 24 calendar days from May 21 to June 13, 2024.

____________________ Orynbasarov A.K.
Personal signature
Date May 20, 2024

Figure 12. Corrected output version of the document

6. CONCLUSION

The proposed solution is to integrate formal languages and neural network technologies into quality management systems. The use of formal languages allows standardizing and clearly structuring regulatory documents, eliminating ambiguities and reducing the probability of errors. Implementation of recurrent neural networks (RNN) provides automated analysis of documents, identification of discrepancies and optimization of management decisions. This combination of technologies helps to improve quality control efficiency, reduce reliance on human error and accelerate regulatory updates, making the quality management system more transparent, adaptive and productive.

The system developed on the basis of the formal language differs from its analogues by implementing a new approach to documenting, which consists in forming the document scheme from special components, which are computer models of structural elements of the document. This process is called "document construction". Using a formal language for normative documents has many advantages. Firstly, it reduces the possibility of different interpretations and interpretations of the document, which avoids various undesirable consequences, such as conflicts, legal errors, lawsuits, etc. Secondly, the formal language simplifies the process of creating and making changes to normative documents, since it has a strict logical structure and format, which makes it possible to automate the process of checking a document for compliance with norms and rules. Thirdly, the use of a formal language of normative documents increases the efficiency of government bodies, since this language allows improving the process of regulating public relations and ensuring more effective protection of citizens' rights and freedoms.

The results of the study emphasize the importance of formal languages and neural networks in industrial automation. Through the use of formal languages, the possibilities of optimizing production processes and harmonizing standards increase. The integration of artificial neural networks increases the efficiency of manufacturing processes and reduces the dependence on the human factor.

Thus, the methods and results presented in this research expand the possibilities of applying formal languages and neural networks to industrial automation. The results of the study contribute to the efficiency of manufacturing systems and harmonization of standards. Further research in this direction should focus on further optimization of manufacturing processes and integration of new technologies.

The implementation of the proposed models faces several challenges, including the inefficient use of work time due to data duplication, the delay in updating and distributing information, limited access to accurate analysis, and the time-consuming process of creating reports and training employees. These issues highlight the need for a more automated decision support system to improve real-time information availability and communication efficiency. The integration of neural network-based systems promises to address these challenges by offering advanced document analysis and interaction capabilities, ultimately enhancing the effectiveness of quality management systems and optimizing production processes.

FUNDING INFORMATION

This research was supported by grants of Ministry of Science and Higher Education of the Republic of Kazakhstan (grant number № AP23488459) “Research and development of scientific-methodological foundations of intellectual system of management of medium and small business entities in Kazakhstan”.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Irbulat Utepbergenov	✓	✓		✓	✓	✓								
Shara Toibayeva		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author Toibayeva Sh.D. on request.

REFERENCES

[1] J. Chang, W. Shen, and J. Yu, “Explaining Kazakhstan’s accession to and compliance with WTO, from a neorealism and a neoliberalism approach,” in *Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)*, 2021, pp. 812–818, doi: 10.2991/assehr.k.211220.139.

[2] M. Olkiewicz, J. Dyczkowska, N. Chamier-Gliszczynski, and T. Królikowski, “Quality management in organizations within the framework of standardized management systems,” *Procedia Computer Science*, vol. 225, pp. 4101–4109, 2023, doi: 10.1016/j.procs.2023.10.406.

[3] D. Anggraini, K. Adi, and J. E. Suseno, “Electronic document management systems implementation across industries: systematic analysis,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 1, p. 264, Oct. 2024, doi: 10.11591/ijeecs.v36.i1.pp264-273.

[4] B. Jose and S. Abraham, “Intelligent processing of unstructured textual data in document based NoSQL databases,” *Materials Today: Proceedings*, vol. 80, pp. 1777–1785, 2023, doi: 10.1016/j.matpr.2021.05.605.

[5] ISO, “The ISO survey,” *ISO*. https://www.iso.org/the-iso-survey.html/ (accessed Aug. 01, 2024).

[6] I. Utepbergenov, S. Toibayeva, and L. Issabekova, “On an integrated automated approach to enterprise quality management,” in *SIST 2023 - 2023 IEEE International Conference on Smart Information Systems and Technologies, Proceedings*, 2023, pp. 299–304, doi: 10.1109/SIST58284.2023.10223500.

[7] A. Jambul and A. Dzhulayeva, “Development of quality management in the republic of Kazakhstan,” *East Asian Journal of Business Management*, vol. 5, no. 2, pp. 23–29, 2015, doi: 10.13106/eajbm.2014.vol5.no2.23.

[8] B. Korabayev, G. Amanova, B. Akimova, K. Saduakassova, and A. Nurgaliyeva, “The model of environmental accounting and auditing as a factor in increasing the efficiency of management decisions at industrial enterprises in the Republic of Kazakhstan,”

Regional Science Policy & Practice, vol. 16, no. 3, p. 12727, Mar. 2024, doi: 10.1111/rsp3.12727.

[9] Y. Zhuman, J. Juman, A. Makulova, B. Kalaganov, A. Tagay, and U. Bastarova, "Mechanism of introduction of quality management system at the enterprises of the construction industry: case of the Republic of Kazakhstan," *Business, Management and Economics Engineering*, vol. 22, no. 02, pp. 240–254, Sep. 2024, doi: 10.3846/bmee.2024.20792.

[10] A. Garrido-Moreno, R. Martín-Rojas, and V. J. García-Morales, "The key role of innovation and organizational resilience in improving business performance: A mixed-methods approach," *International Journal of Information Management*, vol. 77, p. 102777, Aug. 2024, doi: 10.1016/j.ijinfomgt.2024.102777.

[11] N. P. Rana, R. Pillai, B. Sivathanu, and N. Malik, "Assessing the nexus of Generative AI adoption, ethical considerations and organizational performance," *Technovation*, vol. 135, p. 103064, Jul. 2024, doi: 10.1016/j.technovation.2024.103064.

[12] S. Toibayeva and I. Utepbergenov, "Methodology of automated quality management," *Open Engineering*, vol. 14, no. 1, pp. 374–388, Jan. 2024, doi: 10.1515/eng-2022-0545.

[13] D. Makoeva, O. Nagoeva, and I. Gurtueva, "Formal representation of natural language elements in multi-agent system based of self-organization of distributed neurocognitive architectures," *Procedia Computer Science*, vol. 213, pp. 631–635, 2022, doi: 10.1016/j.procs.2022.11.114.

[14] "ScienceDirect. Formal Language – an overview," *ScienceDirect*. www.sciencedirect.com/topics/computer-science/formal-language (accessed Mar. 12, 2023).

[15] A. P. Sergushicheva and E. N. Davydova, "Development and application of a software tool to support the teaching of formal languages," *Open Education*, vol. 25, no. 4, pp. 13–22, Aug. 2021, doi: 10.21686/1818-4243-2021-4-13-22.

[16] D. Jain, M. D. Borah, and A. Biswas, "Summarization of lengthy legal documents via abstractive dataset building: An extract-then-assign approach," *Expert Systems with Applications*, vol. 237, p. 121571, Mar. 2024, doi: 10.1016/j.eswa.2023.121571.

[17] E. Bakinova, A. Basharin, I. Batmanov, K. Lyubort, A. Okhotin, and E. Sazhneva, "Formal languages over GF(2)," *Information and Computation*, vol. 283, p. 104672, Feb. 2022, doi: 10.1016/j.ic.2020.104672.

[18] A. Pettorossi, "Formal grammars and languages," in *Automata Theory and Formal Languages*, Springer International Publishing, 2022, pp. 1–24.

[19] T. Yamakami, "Formal grammars for turn-bounded deterministic context-free languages," in *Theoretical Aspects of Computing – ICTAC 2022*, Springer International Publishing, 2022, pp. 425–441.

[20] R. E. O. Roxas and R. N. C. Recario, "Scientific landscape on opportunities and challenges of large language models and natural language processing," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 1, p. 252, Oct. 2024, doi: 10.11591/ijeecs.v36.i1.pp252-263.

[21] Z. Shi, "Language cognition," in *Intelligence Science*, Elsevier, 2021, pp. 215–266.

[22] R. Quandt and J. Licato, "Problems of autonomous agents following informal, open-textured rules," in *Human-Machine Shared Contexts*, Elsevier, 2020, pp. 221–240.

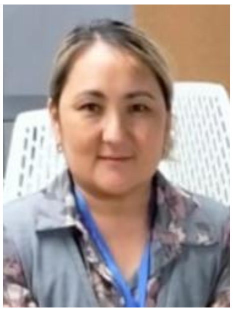
[23] A. Singh, R. M. Parizi, Q. Zhang, K.-K. R. Choo, and A. Dehghantanha, "Blockchain smart contracts formalization: Approaches and challenges to address vulnerabilities," *Computers & Security*, vol. 88, p. 101654, Jan. 2020, doi: 10.1016/j.cose.2019.101654.

[24] K. K. H. Ng, C.-H. Chen, C. K. M. Lee, J. (Roger) Jiao, and Z.-X. Yang, "A systematic literature review on intelligent automation: Aligning concepts from theory, practice, and future perspectives," *Advanced Engineering Informatics*, vol. 47, p. 101246, Jan. 2021, doi: 10.1016/j.aei.2021.101246.

[25] L. Abdykerimova *et al.*, "Analysis of the emotional coloring of text using machine and deep learning methods," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 3, pp. 3055–3063, Jun. 2024, doi: 10.1109/KhPIWeek53812.2021.9570055.

BIOGRAPHIES OF AUTHORS

Irbulat Utepbergenov has a Doctor of Technical Sciences degree, professor in the field of information technologies, graduated from Bauman Moscow State Technical University, specialty "Electronic Computing Machines", Ph.D. in 05.13.06 - Automated Control Systems (1995), professor in the specialty "Informatics, Computer Science and Control". At present she heads the Thematic Research Laboratory "Cyberphysical Systems and Smart Technologies" in Almaty University of Energy and Communication Kazakhstan. Experience of work in the direction of the project - 38 years. Utepbergenov I.T. is the author of 315 scientific works, of which more than 20 publications in the editions indexed in Web of Science and Scopus databases and more than 30 on the subject of the project and 5 copyright certificates for programs and databases. He can be contacted by e-mail: i.utepbergenov@aues.kz.



Shara Toibayeva received a bachelor's degree from Satpayev National Polytechnic University of Kazakhstan and a master's degree in computer science and technology from Turan University (Almaty, Kazakhstan) in 2001 and 2011, respectively, and a Ph.D. degree in automation and control in 2020. Since 2022, he has been an associate professor at the Department of Automation and Control at Gumarbek Daukeev Almaty University of Power Engineering and Communications. Currently, he is an associate professor at the Department of Automation and Control. She can be contacted by e-mail: sh.toibaeva@aues.kz.