Quantum cryptographic systems for protecting digitized data

Lyubomyr Chyrun[1,†], Alona Dorozhynska[2,*,†], and Sergiy Dorozhynskyy[3,†]

[1] *Ivan Franko National University of Lviv, University 1, 79000 Lviv, Ukraine*
[2] *Ukrainian Lingua-Information Fund of NAS of Ukraine, 3, Holosiivskyi avenue, Kyiv, 03039, Ukraine*
[3] *Kyiv University of Intellectual Property and Law, 210, Kharkivske Shosse, Kyiv, 02121, Ukraine*

Abstract

In a world saturated with a huge flow of data, not only the information itself has become important, but also its quality and reliability. Data digitization plays a key role in providing access to true and objective information. Careful processing and use of digitized data is a necessary element, a kind of filter for fighting the spread of misinformation and preserving the credibility of sources among the data flow. The article outlines the range of tasks, approaches and stages of text analysis technology development using dictionaries as an example. The research was conducted for the dictionaries. Based on the conceptual model, the structure of the XML document is built, which is proposed to be used as an intermediary between the paper version of the dictionary and its implementation as an online lexicographic system. In the future, it is planned to build a universal procedure for parsing texts of some fields. The most important stages in terms of information security are the database and website stages, because they are the only ones that can be compromised. That is why reliable security systems, such as quantum cryptographic systems, which are resistant to many types of attacks, should be used at these stages.

Keywords

Digitization, Cyber War, parsing, XML, database, digital space, quantum cryptographic protocols, quantum security system, website.

Introduction

In today's information environment, where the amount of data is growing uncontrollably, understanding and managing the quality and reliability of information becomes a critical task. The huge flow of data coming from various sources affects all areas of our life - from science and business to the social sphere and politics. In this context, data digitization, which consists in converting analog information into a digital format, is a key strategy to ensure effective management of this information flow.

Digitized data provides the possibility of rapid storage, processing and exchange of information. However, their quantity alone does not guarantee quality or truthfulness. Ensuring the accuracy and reliability of data is becoming more important than ever before.

One of the important aspects of using digitized data is its careful processing. Artificial intelligence technologies and machine learning algorithms can be used to detect patterns, anomalies, and verify information. It is important to consider that this process must be accompanied by expert analysis, since the human mind is capable of perceiving context and nuances that may be difficult for algorithms.

In a world where misinformation can become a serious threat, careful filtering and verification of data becomes a necessary element to preserve the credibility of sources and avoid the spread of false information. Digitized data acts as a kind of filter that helps distinguish information from

CLW-2025: Computational Linguistics Workshop at 9th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2025), May 15–16, 2025, Kharkiv, Ukraine
* Corresponding author.
† These authors contributed equally.
Lyubomyr.Chyrun@lnu.edu.ua (L. Chyrun); alonochkatkachyk@mail.com (A. Dorozhynska); dorozhun1706@mail.com (S. Dorozhynskyy)
0000-0002-9448-1751 (L. Chyrun); 0000-0001-6554-6731 (A. Dorozhynska); 0000-0002-5395-6423 (S. Dorozhynskyy)

misinformation, contributing to the creation of an objective and reliable image of the world around us.

Let's emphasize the importance of analyzing dictionaries, "because without them, neither the development of various fields of knowledge, nor modern international communication in the most diverse areas is possible" [4]. However, domestic lexicography still does not have a very rich experience in creating electronic terminological dictionaries.

To conduct the research, we chose the Dictionary of Ukrainian Biological Terminology (SUBT) [1] and the Lexicon, as they have a rich structure.

When analyzing research objects, certain questions and tasks arise.

Among the tasks are the stages of researching the data format and text encoding elements. Building the lexicographic structure of a dictionary and describing the technology of its syntactic analysis is undoubtedly an important issue. It is also important to identify the processes that are the least secure and require serious security solutions. As a result, we should get a generalized secure system of data digitization.

Basic technological stages

The main aspect of data digitization is the formation of a general technology consisting of certain stages and steps. Research occurs in the incrementally converted text on the website. Basic technological stages: paper book, structure of the lexicographic system (L-systems), HTML, text marking with XML tags according to the structure of its L-system (XML-text), conversion of XML-text into a database, website. This process includes steps that can be applied to a variety of paper data, as the following sequence represents an efficient and versatile way to convert texts into digital format.

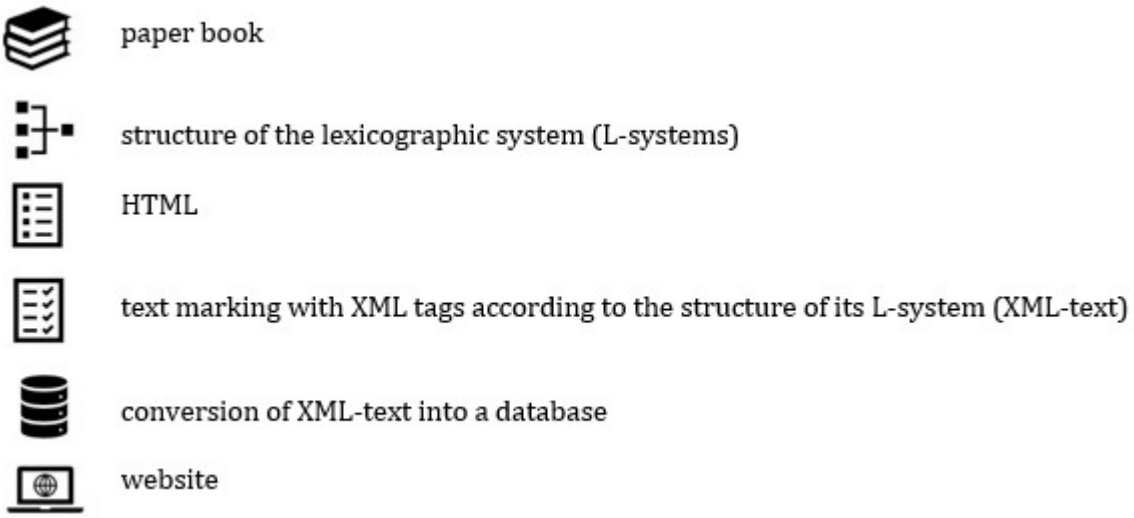


Figure 1: The steps

Consider the individual steps that make up the transformation process.

2.1. Paper book

If the text is accessible in electronic format from the very beginning, then we work with it. Otherwise, the process of scanning, detailed analysis and formatting is added. For convenience, the file can be converted to Word doc format in order to do some text conversions.

For example, we will demonstrate an effective approach for transforming a paper dictionary into an online product "Dictionary of Ukrainian Biological Terminology" (SUBT) "Lexicon of Polish and Ukrainian active phraseology" to an online digital dictionary.

In particular, we had the text of the SBST and the Lexicon from the beginning in the form of a PDF file. For the convenience of further work, these files were converted to Word doc format in order to make some text conversions. Among these transformations, we note the following: the disclosure of abbreviations of a certain type (for example: abrikós → abriko#s and the like.; replacement of stressed letters with a combination of two characters: «літера#». The following replacements were made: á—a#, é—e#, и́—и#, í—i#, ó—o#, ý—y#, ї́—ї#, я́—я#, є́—є#, ю́—ю#, ы́—ы# (for Cyrillic); ý—y# (for Latin). All dictionary articles were processed in this way.

abrikós, -a (ros. abrikós , angl. apricot) Rid derev i kyshiv rodinи rozotsvitих z solodkymy, sokovytymy їstivnymy plodamy, yaki mistit' tsukry, orhanichni kysloty i vitaminy; z nasinня otrymuyut' oliyu. Svitlo i teplolюbni, zharovitrivalі, zasukhostiyki, poshirenі v Yevraziї, SShA.

Figure 2: Text in the dictionary

<B> abriko#s </B>, -a (<I>ros.</I> abriko#s , <I>angl.</I> apricot) Rid derev i kyshiv rodinи rozotsvitих z solodkymy, sokovytymy їstivnymy plodamy, yaki mistit' tsukry, orhanichni kysloty i vitaminy; z nasinня otrymuyut' oliyu. Svitlo i teplolюbni, zharovitrivalі, zasukhostiyki, poshirenі v Yevraziї, SShA.

Figure 3: Text after processing

2.2. Construction of a conceptual model of the lexicographic system

In accordance with the theory of lexicographic systems [2], the general structure of the dictionary article of the DUBT is revealed. The structure of the L-system is presented in the form:

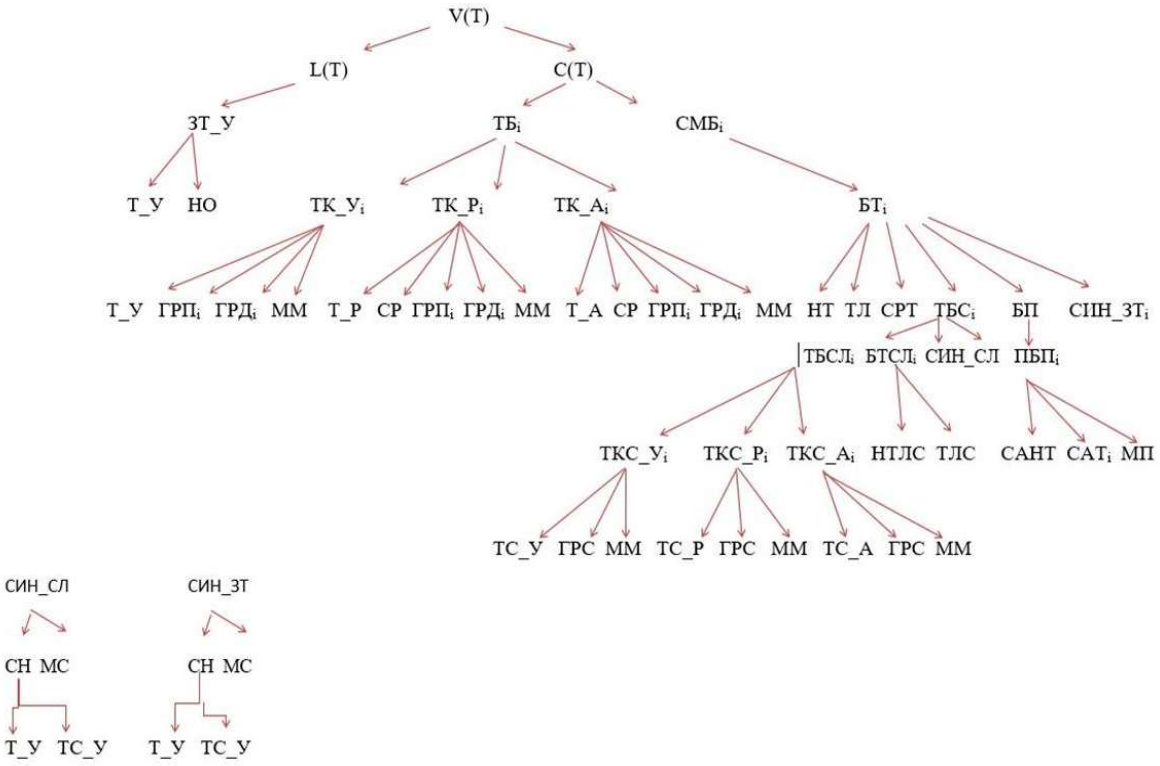


Figure 4: General scheme of the dictionary article of DUBT

The following notation is used in this scheme:

СС — dictionary article
ЗТ_У — the title term is Ukrainian
ТБi — terminological block
ТК_Уi — terminological complex ukr.
Т_У — Ukrainian term
ГР— grammatical trailer

НО — homonym number
ММ — language marker (ukr)
ТК_Рі — terminological complex rus.
Т_Р — Russian term
ГР — grammatical trailer
ММ — language marker (rus)
СР — semantic trailer
ТК_Аі — terminological complex eng.
Т_А — English term
ГР — grammatical trailer
ММ — language marker (eng)
СР — semantic trailer
СМБі — semantic block
БТі — interpretation block
НТ — interpretation number
ТЛІ — interpretation
СРТ — semantic trailer to ТЛІ
БТСі — block of terminol. phrases
ТБСЛІі — terminol. block of phrases
ТСК_Уі — terminological complex of phrases ukr.
ТС_У — terminological phrase ukr.
ГРС — grammatical remark of the phrase
ММ — language marker (укр)
ТСК_Рі — terminological complex of phrases rus.
ТС_Р — terminological phrase rus.
ММ — language marker (rus.)
ТСК_Аі — terminological complex of phrases eng.
ТС_А — terminological phrase eng.
ММ — language marker (eng)
БТСЛІі — block of interp. of phrases
ТЛІС — interpretation
НТЛІС — number ТЛІС
СИНі — synonymous block
СН — synonym
МС — synonym marker (син.)

БП — link block
ПБПі — subblock of links
САНТ — sender
САТі — recipient
МП — link marker (див.)

Example 1

adventy#vnyi (ros. adventy#vnyy, anhl. adventive) 1.Yakyi rozvyvaiet′sia ne z embrional′nykh tkanyn tochky rostu, a iz starishykh chastyn roslyny; adventy#vna bry#n′ka dıv. bry#n′ka: bry#n′ka adventy#vna; adventy#vna embrioni#ia dıv. embrioni#ia: embrioni#ia adventy#vna; adventy#vna poliembrioni#ia dıv. poliembrioni#ia: poliembrioni#ia nutseli#arna [adventy#vna]; adventy#vnyi za#rodok dıv. za#rodok: za#rodok adventy#vnyi; adventy#vnyi o#rhan dıv. o#rhan: o#rhan adventy#vnyi; adventy#vnyi pa#hin dıv. pa#hin: pa#hin adventy#vnyi 2. Zanesena liudynoiu roslyna v tu mistsevist′, de vona rani she ne rosla; adventy#vna rosly#na dıv. rosly#na: rosly#na adventy#vna; adventy#vna vyd dıv. vyd: vyd adventy#vnyi.

According to the structure of the dictionary article, we will highlight all the structural elements

ТБ [terminolohichnyi blok]: adventy#vnyi (ros. adventy#vnyi, anhl. adventive)
ТК_У [terminolohichnyi kompleks ukrains′kyi]: adventy#vnyi
ЗТ [zaholovnyi termin]: adventy#vnyi
ТК_Р [terminolohichnyi kompleks rosiis′kyi]: ros. adventy#vnyi
ММ [marker movy]: ros.
Т_Р [termin rosiis′kyi]: adventy#vnyi
ТК_А [terminolohichnyi kompleks anhliis′kyi]: anhl. adventive
ММ [marker movy]: anhl.
Т_А [termin anhliis′kyi]: adventive

СМБ [semantychnyi blok]: 1. Yakyi rozvyvaiet′sia ne z embrional′nykh tkanyn tochky rostu, a iz starishykh chastyn roslyny; adventy#vna bry#n′ka dyv. bry#n′ka: bry#n′ka adventy#vna; adventy#vna embrioni#ia dyv. embrioni#ia: embrioni#ia adventy#vna; adventy#vna poliembrioni#ia dyv. poliembrioni#ia: poliembrioni#ia nutselia#rna [adventy#vna]; adventy#vnyi za#rodok dyv. za#rodok: za#rodok adventy#vnyi; adventy#vnyi o#rhan dyv. o#rhan: o#rhan adventy#vnyi; adventy#vnyi pa#hin dyv. pa#hin: pa#hin adventy#vnyi

2. Zanesena liudynoiu roslyna v tu mistsevist′, de vona ranishe ne rosla; adventy#vna rosly#na dyv. rosly#na: rosly#na adventy#vna; adventy#vna vyd dyv. vyd: vyd adventy#vnyi.

БТ1 [blok tlumachen′]: 1. Yakyi rozvyvaiet′sia ne z embrional′nykh tkanyn tochky rostu, a iz starishykh chastyn roslyny; НТ [nomer tlumachennia]: ТЛ1 [tlumachennia]: Yakyi rozvyvaiet′sia ne z embrional′nykh tkanyn tochky rostu, a iz starishykh chastyn roslyny;

БТ2 [blok tlumachen′]: 2. Zanesena liudynoiu roslyna v tu mistsevist′, de vona ranishe ne rosla; НТ [nomer tlumachennia]: ТЛ2 [tlumachennia]: Zanesena liudynoiu roslyna v tu mistsevist′, de vona ranishe ne rosla;

БП1 [blok posylan′]: adventy#vna bry#n′ka dyv. bry#n′ka: bry#n′ka adventy#vna; adventy#vna embrioni#ia dyv. embrioni#ia: embrioni#ia adventy#vna; adventy#vna poliembrioni#ia dyv. poliembrioni#ia: poliembrioni#ia nutselia#rna [adventy#vna]; adventy#vnyi za#rodok dyv. za#rodok: za#rodok adventy#vnyi; adventy#vnyi o#rhan dyv. o#rhan: o#rhan adventy#vnyi; adventy#vnyi pa#hin dyv. pa#hin: pa#hin adventy#vnyi
ПБП1 [pidblok posylan′]: adventy#vna bry#n′ka dyv. bry#n′ka: bry#n′ka adventy#vna;
ПБП2 [pidblok posylan′]: adventy#vna embrioni#ia dyv. embrioni#ia: embrioni#ia adventy#vna;
ПБП3 [pidblok posylan′]: adventy#vna poliembrioni#ia dyv. poliembrioni#ia: poliembrioni#ia nutselia#rna [adventy#vna];
ПБП4 [pidblok posylan′]: adventy#vnyi za#rodok dyv. za#rodok: za#rodok adventy#vnyi;
ПБП5 [pidblok posylan′]: adventy#vnyi o#rhan dyv. o#rhan: o#rhan adventy#vnyi;
ПБП6 [pidblok posylan′]: adventy#vnyi pa#hin dyv. pa#hin: pa#hin adventy#vnyi
БП2 [blok posylan′]: adventy#vna rosly#na dyv. rosly#na: rosly#na adventy#vna; adventy#vna vyd dyv. vyd: vyd adventy#vnyi.
ПБП1 [pidblok posylan′]: adventy#vna rosly#na dyv. rosly#na: rosly#na adventy#vna;
ПБП2 [pidblok posylan′]: adventy#vna vyd dyv. vyd: vyd adventy#vnyi.

Each sub-block consists of an addressee1, a link marker and an addressee2.
ПБП2 [pidblok posylan′]: adventy#vna vyd dyv. vyd: vyd adventy#vnyi.
САНТ [adresant]: adventy#vna vyd
МП [marker posylan′]: dyv.
САТ [adresat]: vyd: vyd adventy#vnyi.

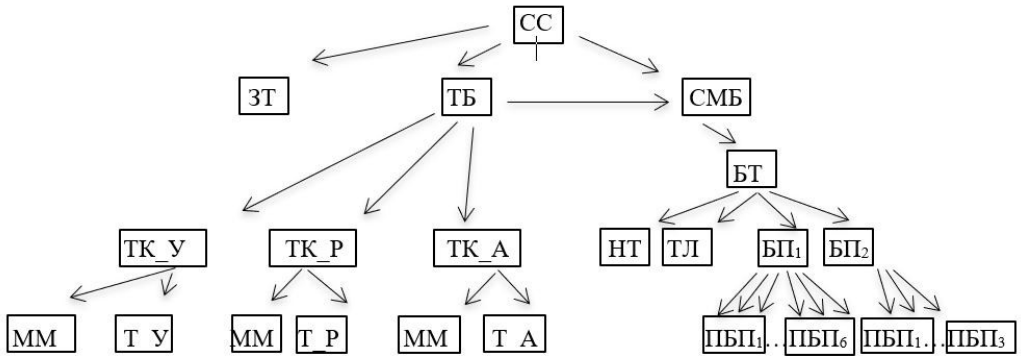


Figure 5: Outline of a dictionary article

2.3. HTML

The next stage is the markup of the main HTML tags. This will further simplify the algorithm of the program that forms the XML. All further work is based on XML file processing.

The example shows the markup for dictionari. A container was immediately developed for the Lexicon, as its structure is simple. And some steps in the work can be skipped.

Example 2

<B>a#apa</B> (<I>ros.</I> aa#pa, <I>angl.</I> aapa) 1. Bezlisyi, duzhe obvodnenyi, hriadkovo-mocharnyi vyd bolota na pivnochi Yevrazii.

<B>aba#ka</B> (<I>ros.</I> aba#ka, <I>angl.</I> abacus) 1. Bahatorichna trav'ianysta roslyna rodyny bananovykh, iaku kultyvuiut′ u tropikakh; vykorystovuiut′ u narodnomu hospodarstvi dlia vyhotovlennia kanativ, motuzok, korabel′nykh snastei toshcho.

<B>abaksi#al'nyi</B> (<I>ros.</I> abaksia#l'nyĭ, <I>angl.</I> abaxial) 1. Yakyi stosuietsia nyzhnoi poverkhni lystovoho orhana.

<B>abera#ntnyi</B> (<I>ros.</I> aberra#ntnyĭ, <I>angl.</I> aberrational) 1. Nezvychainyi, shcho vidkhylyvsia vid normy. Aberantni pryklady parazytnykh vidnosyn predstavleni hiperparazytyzmom.

2.4. Marking text with XML tags according to its structure L-systems

The XML document provides an opportunity to explain all the structural elements we have defined and the relationships between them. To ensure automatic marking of the dictionary text with XML tags, a program was developed that separates the elements of the text structure in accordance with the structure of the L-system and HTML markup.

According to the general CML structure of the article, which was considered in the previous articles, it is possible to show the analysis of the article with an example.

Example 3

<CC>

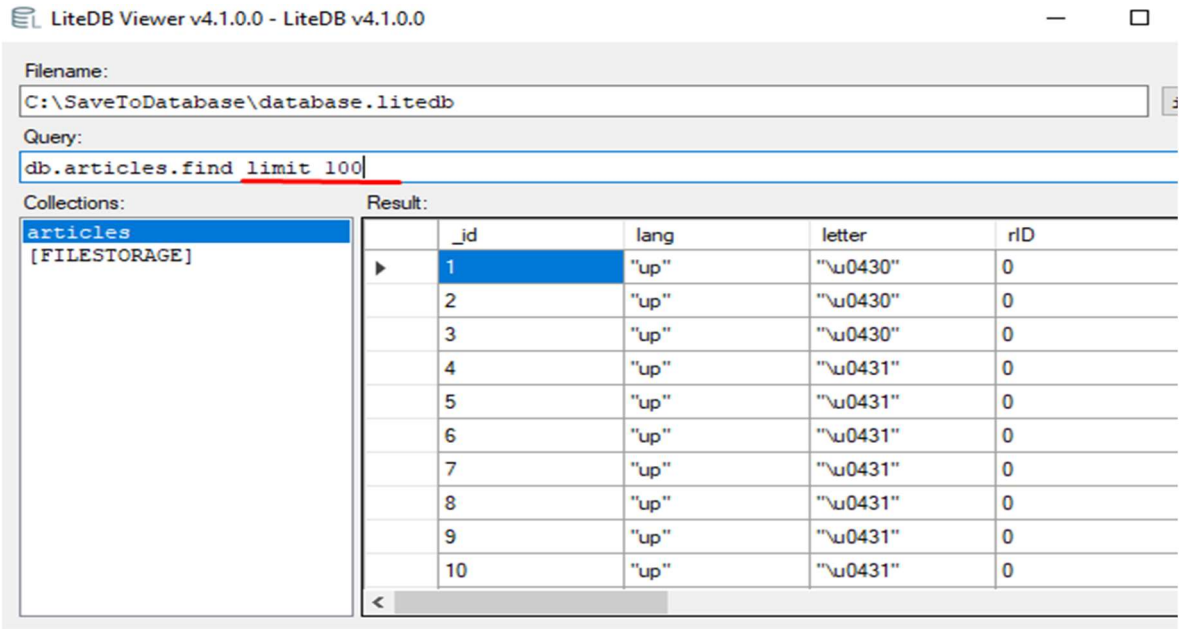
<текст_СС><![CDATA[<B>adventy#vnyi</B> (<I>ros.</I> adventy#vnyĭ, <I>angl.</I> adventive) 1. Yakyi rozvyvaiet′sia ne z embrional′nykh tkanyn tochky rostu, a iz starishykh chastyn roslyny; <B>adventy#vna bry#n′ka</B> <I>dyv.</I> <B>bry#n′ka: bry#n′ka adventy#vna</B>; <B>adventy#vna embrioni#ia</B> <I>dyv.</I> <B>embrioni#ia: embrioni#ia adventy#vna</B>; <B>adventy#vna poliembrioni#ia</B> <I>dyv.</I> <B>poliembrioni#ia: poliembrioni#ia nutselia#rna [adventy#vna]</B>; <B>adventy#vna za#rodok</B> <I>dyv.</I> <B>za#rodok: za#rodok adventy#vnyi</B>; <B>adventy#vnyi o#rhan</B> <I>dyv.</I> <B>o#rhan: o#rhan adventy#vnyi</B>; <B>adventy#vnyi pa#hin</B> <I>dyv.</I> <B>pa#hin: pa#hin adventy#vnyi</B> 2. Zanesena liudynoiu roslyna v tu mistsevist′, de vona ranishe ne rosla; <B>adventy#vna rosly#na</B> <I>dyv.</I> <B>rosly#na: rosly#na adventy#vna</B>; <B>adventy#vna vyd</B> <I>dyv.</I> <B>vyd: vyd adventy#vnyi</B>.]]></текст_СС>

<ЗТ>adventy#vnyi</ЗТ>
<ТБ nomer="1">
<ТК_У nomer="1">
<Т_У>adventy#vnyi</Т_У>
<ММ>ukr.</ММ>
</ТК_У>
<ТК_Р nomer="1">
<Т_Р>adventy#vnyĭ</Т_Р>
<ММ>ros.</ММ>
</ТК_Р>
<ТК_А nomer="1">
<Т_А>adventive</Т_А>
<ММ>angl.</ММ>
</ТК_А>
</ТБ>
<СМБ nomer="1">
<БТ nomer="1">
<НТ>1</НТ>

<TЛ>Yakyi rozvyvaiet′sia ne z embrional′nykh tkanyn tochky rostu, a iz starishykh chastyn roslyny</TЛ>
</БТ>
<БТ nomer="2">
<НТ>2</НТ>
<TЛ>Zanesena liudynoiu roslyna v tu mistsevist′, de vona ranishe ne rosla</TЛ>
</БТ>
<БП nomer="1">
<САНТ>adventy#vna bry#n′ka</САНТ>
<САТ>bry#n′ka: bry#n′ka adventy#vna</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="2">
<САНТ>adventy#vna embrioni#ia</САНТ>
<САТ>embrioni#ia: embrioni#ia adventy#vna</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="3">
<САНТ>adventy#vna poliembrioni#ia</САНТ>
<САТ>poliembrioni#ia: poliembrioni#ia nutselia#rna [adventy#vna]</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="4">
<САНТ>adventy#vna za#rodok</САНТ>
<САТ>za#rodok: za#rodok adventy#vnyi</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="5">
<САНТ>adventy#vnyi o#rhan</САНТ>
<САТ>o#rhan: o#rhan adventy#vnyi</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="6">
<САНТ>adventy#vnyi pa#hin</САНТ>
<САТ>pa#hin: pa#hin adventy#vnyi</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="7">
<САНТ>adventy#vna rosly#na</САНТ>
<САТ>rosly#na: rosly#na adventy#vna</САТ>
<МП>dyv.</МП>
</БП>
<БП nomer="8">
<САНТ>adventy#vna vyd</САНТ>
<САТ>vyd: vyd adventy#vnyi</САТ>
<МП>dyv.</МП>
</БП>
</СМБ>
</СС>

2.5. Conversion of XML-text into a database

The creation of a lexicographic database is based on the processing of a strict and meta-information-rich structure of the input XML file. Sequential traversal of all nodes of the hierarchical structure of the XML-representation of dictionary articles made it possible to linearly translate them into the form of software objects - instances of the class of dictionary articles, which are stored in an explicit form in the LiteDB document type database. During this process, additional parameters for their characterization were selected from the structurally marked text of the articles, which made it possible to create an expanded index of register units and corresponding explanatory parts of this dictionary.



_id	lang	letter	rID
1	"up"	"\u0430"	0
2	"up"	"\u0430"	0
3	"up"	"\u0430"	0
4	"up"	"\u0431"	0
5	"up"	"\u0431"	0
6	"up"	"\u0431"	0
7	"up"	"\u0431"	0
8	"up"	"\u0431"	0
9	"up"	"\u0431"	0
10	"up"	"\u0431"	0

Figure 5: A fragment of the Lexicon database

3. Quantum-cryptographic systems for protecting digitized data

Protecting digitized data is a fundamental aspect of modern information management. In a world where vast amounts of information are stored and processed electronically, the security of this data is becoming critical. Digitized information is not only a key resource for businesses and organizations, but also contains sensitive data such as personal customer information, financial data, medical records, etc.

Data protection involves a number of aspects, such as ensuring confidentiality, preventing unauthorized changes (ensuring integrity), and ensuring that data is available to legitimate users. With cybercrime and information leakage risks constantly evolving, the importance of data protection is manifested in economic losses, legal consequences, loss of consumer confidence, and preservation of organizations' reputation. In addition, regulatory requirements and statutory laws oblige many companies to comply with specific standards for the protection of personal data and confidential information. Non-compliance can result in severe sanctions and fines.

The use of quantum key distribution (QKD) protocols in cryptography is determined by the need to provide a higher level of security and resistance in the field of key exchange and protection of confidential information. Quantum cryptography uses the principles of quantum mechanics to generate and exchange keys, ensuring absolute confidentiality due to the principle of uncertainty of quantum states. QKD protocols include methods for detecting any attacks or interception attempts, making them particularly effective in the face of cyber threats and the growing interest in quantum computing. These protocols not only provide a response to current cybersecurity challenges, but

also take into account the possible development of quantum technologies, such as quantum computers, which may threaten traditional cryptographic methods. The use of QKD also helps to ensure security at the physical level of the network, where ensuring that keys are inaccessible to attackers becomes a key factor.

The Twin Field protocol is an important tool in the field of protecting digitized data through the use of quantum cryptography. If important improvements are applied, such as quantum channel multiplexing and quantum identification, significant improvements in the security of quantum key exchange can be achieved.

The total number of quantum states required to multiplex N channels can be calculated by the formula:

$$Nt = 2N(M + L),$$

where N_t is the total number of quantum states required to multiplex all N channels. When multiplexing quantum states from N channels, the total quantum state is transmitted to channel Z:

$$\left|\varphi_Z\right\rangle = \bigotimes_{j=1}^{N} \left|\varphi_j^{(Z)}\right\rangle,$$

where $\varphi_j^{(Z)}$ is the quantum state transmitted in the j-th quantum channel using channel Z. Similarly, the total quantum state is transmitted for channel X:

$$\left|\varphi_X\right\rangle = \bigotimes_{j=1}^{N} \left|\varphi_j^{(X)}\right\rangle,$$

where $\varphi_j^{(X)}$ is the quantum state transmitted in the j-th quantum channel via channel X. The formula for multiplexing quantum channels:

$$\left|\psi\right\rangle_{ABCD} = \frac{1}{\sqrt{2}} \left(\left|0\right\rangle_A \left|\psi_0\right\rangle_B \left|\psi_0\right\rangle_C + \left|1\right\rangle_A \left|\psi_1\right\rangle_B \left|\psi_1\right\rangle_C\right).$$

The use of quantum channel multiplexing contributes to the efficient transmission of information, increasing bandwidth and optimizing the use of resources. This is especially true in large networks where data transmission speed and reliability are important. Quantum identification, using the principles of quantum mechanics, provides authentication of communication partners, making it impossible to be monitored or attacked by illegitimate parties. This strengthens security and trust in the exchange of quantum keys. Reducing the vulnerability to quantum espionage that can arise from the physical interception of quantum bits is an important aspect of Twin Field QKD security. This makes it an effective tool to protect against attacks involving the use of quantum technologies.

The Twin Field protocol in the context of forming a database for digitized data includes several critical steps to ensure a high level of security and confidentiality. The initial step is to generate a quantum key using the principles of quantum cryptography. During this process, a combination of quantum particles determines a unique key that will be used to encrypt and decrypt the digitized data. The resulting quantum key is used to encrypt the digitized data before it is transmitted over open communication channels. It is important to note that even if the channels or the data itself is intercepted, it remains encrypted and invaluable to unauthorized users without the appropriate quantum key. One of the key aspects of the protocol is the use of quantum identification, which provides authentication and authorization of communication partners. This makes it impossible to be monitored or attacked by illegitimate parties, building trust in the individuals or systems exchanging data. The quantum key used to protect the digitized data can be periodically updated to further reduce the risk of hacking and increase system resilience. The storage and management of these keys is a critical element, as the security of the entire system is largely determined by the secrecy of the quantum key.

Table 1
Characteristics of the protection formation process

Process Stages	Reliability	Speed	Cost	Security	Implementation Complexity
Quantum Key Generation	High	High	Moderate	High	High
Data Encryption	High	Moderate	Low	High	Moderate
Encrypted Data Transmission	High	High	Moderate	High	Moderate
Decryption and Decoding	High	Moderate	Low	High	Moderate
Key Update (optional)	Moderate	Low	Moderate	High	Moderate
Key Storage and Management	High	Moderate	Moderate	High	Moderate

Acknowledgements

Digitized data, by its very nature, helps reduce misinformation through several mechanisms. First, they provide the ability to verify information through metadata such as source and time of creation, which provides the ability to verify its authenticity. However, they can be analyzed with the help of various algorithms and artificial intelligence tools, which allows detecting possible anomalies and changes in information that may indicate disinformation. The main criterion is that digital data is available to a wide range of users, which, due to compliance with trusted sources, contributes to the dissemination of truthful information. In addition, they provide an opportunity to compare data and conduct data analysis to identify inconsistencies and refute false information facts.

The presented stages are universal and can be applied to many types of paper publications, because each of them has its own structure. An important process will be to correctly determine the type of publication and build a conceptual model of the L-system, after which the technology will be applied.

In the future, it is planned to improve the basic technological stages in order to generalize the transformation of texts of various types. Also develop XML structures (containers) for certain industries and types of data.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] D. M. Grodzinsky, L. O. Simonenko and other. Ukrainian biological terminology Dictionary. – K.: KMM, 2012. 746 p.

[2] Linguistic and information studies: works of the Ukrainian language and information fund of the NAS of Ukraine: in 5 volumes / V. A. Shyrokov et al. Vol. 1: Kyiv. 2018. 271 p.

[3] Shyrokov V. A. (Eds.) (2011). Computer lexicography. Kyiv.: Science opinion.

[4] Olga Karpova, Faina Kartashkova Lexicography and Terminology: A Worldwide Outlook. Cambridge : Cambridge Scholars Publishing, 2009. 205 p.

[5] Shyrokov V. A. etc. Linguistic and information studies: works of the Ukrainian Language and Information Fund NAS of Ukraine: in 5 V. V. 5 : Virtualization of linguistic technologies. Kyiv. Ukrainian Lingua-Information Fund of NAS of Ukraine. 2018. 239 p. URL:

https://movoznavstvo.org.ua/files/Ling_inf_studio_TOM_5_umif_B5.pdf. doi: 10.33190/978-966-02-8683-2/8690-0

[6] I. Kernerman A multilingual trilogy: Developing three multi-language lexicographic datasets. Electronic Lexicography in the 21st Century: Linking lexical data in the digital age. Proceedings of eLex 2015, 11-13 August 2015. 372-383p. URL: https://elex.link/elex2015/