Cognitive ontological models of cyber security in the social networks

A.Aktayeva
Abay Myrzakhmetov Kokshetau University,
Kokshetau, Kazakhstan
aaktaewa@list.ru

R.Niyazova
L.N.Gumilyov Eurasian National University,
Astana,Kazakhstan
rozamgul@list.ru

E.Makatov
Abay Myrzakhmetov Kokshetau University,
Kokshetau, Kazakhstan
er_han89@mail.ru

A.Dautov
Abay Myrzakhmetov Kokshetau University,
Kokshetau, Kazakhstan
d.abeke@mail.ru

***Abstract*—The structure and basic principles of technology for increasing the probability of identifying subjects of information processes of open Internet resources based on ontology methods are considered. Based on this ontology the knowledge base intended for creation of the program systems supporting ensuring information security has been realized. The developed ontological knowledge base has been used when developing the software complex intended for identification of the user of social networks when ensuring information security, monitoring and preventing threats. This article is next in a series of articles by the authors in which they continue to monitor and analyse the current state and new tendencies in the field of information security and safety of information.**

***Keywords*—ontology, knowledge base, information security, Social network, SPARQL, identification**

I. Introduction

During rapid growth in application of information technologies, the successful decision of problems in the field of an information security and protection of the information assumes the more effective activity during a safety in all areas of ability to live of the person. Safety, reliability and privacy are also a part of cyber security. When these systems begin to have a direct physical impact, society has now become responsible for the safety of people and environments.

Historically, the term cybersecurity referred to all the technologies associated with the gathering, processing, storing, and security of information.

The rapid increase in computing and communications power has raised considerable concern about privacy both in the public and private sector. However, with the passage of time and the progress of technologies, the term has acquired different connotations. High latency and frequently the international character of such crimes raise their public danger of the world community.

The modern term, information security or cybersecurity, came into widespread use only in the late 1970s and is now used generally to embrace both computer and communication technologies and their common basis protection – microelectronic technology and all the related software technology. The term Cyberspace notionally represents the various environments that have evolved to support networked computing across the globe. Cybersecurity builds on traditional information security to deal with the evolution of Cyberspace as it grows to include very large and complex systems, mobile computing platforms, cloud computing platforms, and an array of sensors and actuators. The transnational and transboundary character of many products ICT and the international coherence of social networks are used cybersecurity with a view of fulfillment of illegal actions concerning users and the owners the Internet-resources placed in a transnational segment, as well as an AIS, cooperating with the Global network.

The situation is aggravated with the stereotypes which have taken root in a world society about impunity so-called «cybersecurity», the uselessness of accepted measures on strengthening the area of safe use ICTS (info-communication technologies and systems), the limited possibilities of a society on attraction to the responsibility of hi-tech crimes guilty of fulfillment, despite of the developed legal information institutions security in the field of social networks [18].

Cybersecurity should be considered as a sustainable state of the information sphere, ensuring its integrity and protection of ICT infrastructure facilities in the presence of adverse internal and external influences because of awareness of the society of its values, vital interests and development goals.

The neglect of cybersecurity policy when using the resources of social networks of the Internet leads to an increased risk for privacy, unauthorized use or modification of publicly available personal data, as well as disclosure of users' personal data or their transnational accessibility to criminal communities or intelligence agencies of different countries. This, in turn, causes the need to control the subjects of information processes to identify possible areas of information impact and impact on users of social networks on the Internet. Within the framework of this task, it is extremely important to identify the subjects of information processes that can legally distribute 'unreliable or contradictory' messages [15, 16]. Many Internet resources and services, such as forums, portals (social networking resources), online stores, face various manifestations of problems of manipulation and artificial formation of public opinion, by 'organizing' focused thematic dialogues in which many users have multiple account accounts. The possibility of using social portals for information dissemination and insufficient functionality of authentication and authentication mechanisms for users who leave messages determines a few directions for improving protection systems and information security monitoring systems of ICTS.

In this connection, the problem arises of increasing the probabilistic indicators of the quality of methods for identifying users of various Internet portals. One of the promising areas of research in this area is the modeling of cybersecurity systems using the principles of cognitive ontological representation of knowledge, considering the specifications of this subject area.

II. Principles of cognitive ontological representation of knowledge

Ontologies were proposed for the declarative representation of knowledge and are defined in general terms as a special kind of knowledge base or as a 'specification of conceptualization' of any subject area. This means that in the subject area, based on the classification of the basic terms, the main concepts - concepts are singled out, and the connections between them are established - conceptualization. Then the ontology can be represented graphically or described in one of the formal languages (formal ontology) is the ontology specification process.

The ontological representation of knowledge is used for the semantic integration of information resources, adequate interpretation of the content of text documents presented in natural language [15,16,19]. The basis for developing the principles of cognitive ontological representation of knowledge is a diagram reflecting the interrelationships of the basic concepts of security, given in the International Standard for Cybersecurity ISO / IEC 27032: 2012. The ontology formed on this basis reflects only the concepts described in this standard, and only partial details the various aspects that need to be considered when designing a system for ensuring cybersecurity of the social network infrastructure of the Internet [15,16,19]. Then, as can be seen from Figure 1, to ensure the safety of the social networking infrastructure of the Internet, it is necessary to take into account a variety of the factors reflecting the characteristics of all stakeholders, their resources, possible threats and take appropriate response measures against adverse internal and external impacts on ICT infrastructure facilities. Cognitive models are used to model information security threats, and event models are used to model development options for different situations (see Fig.2). And cognitive modeling of ontology is the construction of cognitive models (oriented graphs) in which vertices correspond to concepts, and arcs to connections between the factors [15, 16, 19].

Event-based modeling – the construction of behavioral models (the behavior of users of social networks), and as objects of modeling can be considered as people and technical objects. The essence of the event modeling method is to track the sequence of events on the model in the same order as they would occur in the real system [15,16,19]. The joint use of cognitive and event modeling allows obtaining a more objective assessment of the situation in social networks. For this we introduce ontologies of events used in the transition from cognitive to event models.

When formalizing the ontology problem, it is important to note that: the ontology is one of the tools needed to model the domain; ontology contains a list of key concepts of the given subject area and specification of their meaning; knowledge of the meaning of key concepts represented by the ontology should be obvious to any expert in the given subject area, knowledge bases are developed on the basis of ontologies [12]. The proposed concept supports the technology of researching the directions of the development of the social Internet infrastructure considering the information security requirements. We assume that the infrastructure of social networks of the Internet is defined as

$VSN = \{O, E, MC, MS\} \cup TSN,$
where O is the set of ontologies;
E – set of descriptions of use cases;
MC – a set of cognitive models;
MS is the set of event models;
TSN – an ICT / social networking support tool that includes a description of the knowledge presented in the form of ontologies, descriptions of precedents, cognitive and event models, and means of operating them.

The formal ontology of the subject domain $S\delta$ is the pair S and δ, where δ is the set of key notions of the domain, and S is the set of analytic sentences describing the meaning of these key concepts [2,15,16].

As a result of the analysis of the state of the subject area of the identification of users of the social Internet, it is necessary to distinguish the following: – due to the wide possibilities for providing anonymity to user's social networks on the Internet, the identification methods are of particular importance. However, this method does not take into account changes in the technical characteristics of the device - methods for determining the authorship of the text used by classical linguists show good results for large volumes of text that have undergone correction, but require thorough adaptation for the processing of short messages – to improve the quality of methods for identifying users of social It is necessary to develop a cortege of linguistic signs of a short message, allowing to take into account the peculiarities of construction and of the identifiers.

To date, the most popular methods of identification, using technical characteristics, primarily, such as:

- HTTP Cookie;
- IP-address;
- MAC-address;
- geolocation data;
- data about the operating system, browser, hardware parameters (resolution and screen size, CPU, etc.).

Identification methods using technical characteristics of IKTS are effective for searching for 'single trolls' or unscrupu-

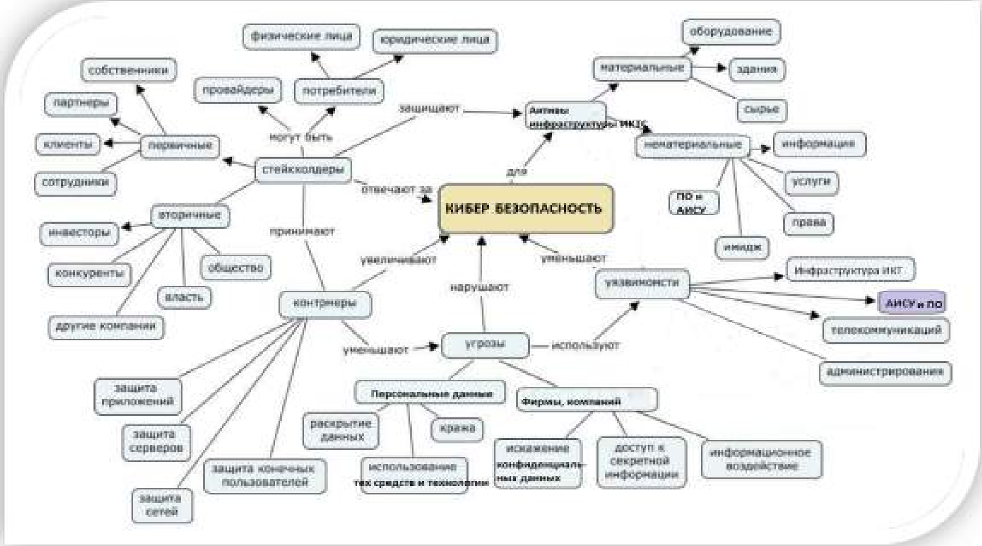


Figure 1. Ontology of Cybersecurity for ISO / IEC 27032: 2012 [15,16].

lous users, but are ineffective for combating and identifying organized astroturfing (AstroTurf) conducted by special organizations that can provide a change in these characteristics [2,4,5,12].

All these features must be considered in the identification process to improve the quality of the results obtained. Thus, based on the received model of a text message containing information about the lexical, graphmatic and syntactic components, it becomes possible to define a profile for each user of the Internet portal.

A user profile is a collection of data and settings from the user's environment. The construction of a user profile is possible on the basis of a number of technical characteristics and statistical data. This approach to creating a profile cannot always give a reliable result. The proposed user profile is especially important in cases where it is possible to replace, clone several technical characteristics of devices, i.e. almost unambiguous identification of the user is impossible [2,4,5,12].

The method of creating an Internet user profile involves the implementation of a number of steps:

- Processing of user messages within the Internet portal;
- Parsing messages by parts of speech followed by the use of templates to highlight the most common constructions.

Lexicographic analysis of the message and allocation of structures in accordance with the patterns described:

- Statistics on the use of punctuation marks and special symbols;
- Selection of lexical constructions based on words and word forms of the language, as well as the identification of thematic special words and phrases specific to a audience online.

The implementation of the proposed method for constructing the user profile of the Internet portal is aimed at solving the indicated tasks in cases where several people use the same PC, or the messages are left by users located on the same local subnet. Figure 3 shows the process of creating a user profile using linguistic characteristics [2,4,5,12].

III. Applications of the ontological representation of knowledge for the construction of the user profile in social network

The ontological approach to the representation of knowledge makes it possible to apply the existing and tested approvals of the advanced profile analytical requests for each user of the Internet portal. To use the ontological representation of knowledge to build a user profile of the Internet portal and to compile a cognitive model of the domain, it is necessary to:

1) Identify significant factors;
2) Construct a matrix of mutual influences;
3) Determine the initial trends of changing factors.

Thus, the infrastructure of the social networks of the Internet includes a knowledge space that integrates: ontological models of knowledge in the field of IS research, knowledge base about precedents in social networks and knowledge bases containing cognitive models of strategic threats of information security and event models of development and consequences of events in social networks, and also tools for describing knowledge (see Fig. 3).

Information extraction is traditionally aimed at finding information that describes a certain area of knowledge specified by the data structure. Ontologies are just a formal domain model expressed, for example, in the form of a graph of concepts and

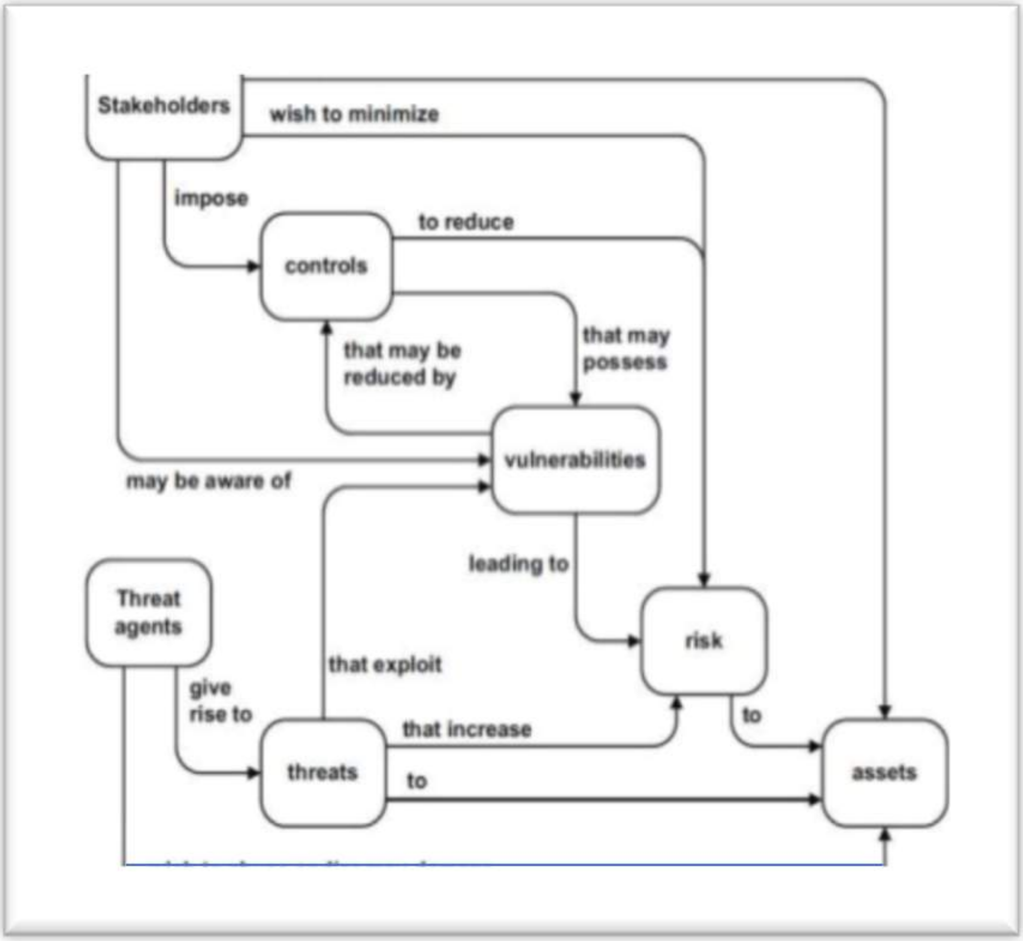
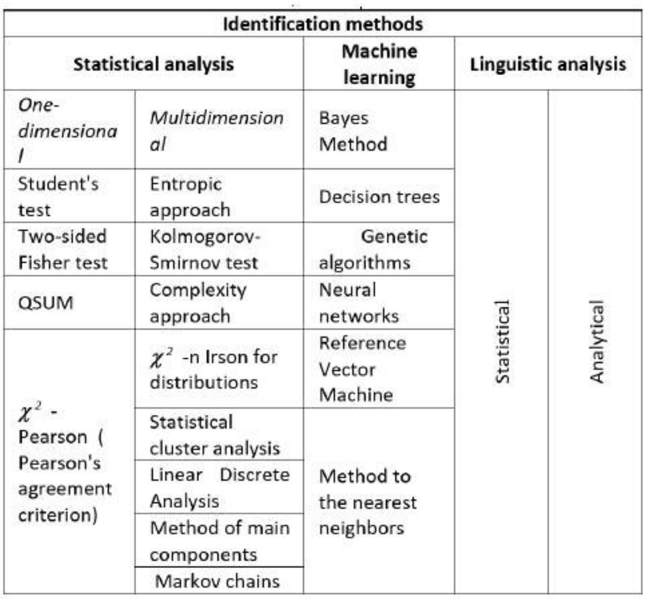


Figure 2. Basic concepts of security and the nature of the relationship between them [21].



Identification methods				
Statistical analysis		Machine learning	Linguistic analysis	
One-dimensiona l	*Multidimension al*	Bayes Method	Statistical	Analytical
Student's test	Entropic approach	Decision trees		
Two-sided Fisher test	Kolmogorov-Smirnov test	Genetic algorithms		
QSUM	Complexity approach	Neural networks		
χ^2 - Pearson (Pearson's agreement criterion)	χ^2 -n Irson for distributions	Reference Vector Machine		
	Statistical cluster analysis	Method to the nearest neighbors		
	Linear Discrete Analysis			
	Method of main components			
	Markov chains			

Figure 3. Basic Methods of Identifying the Authors of Messages (posts).

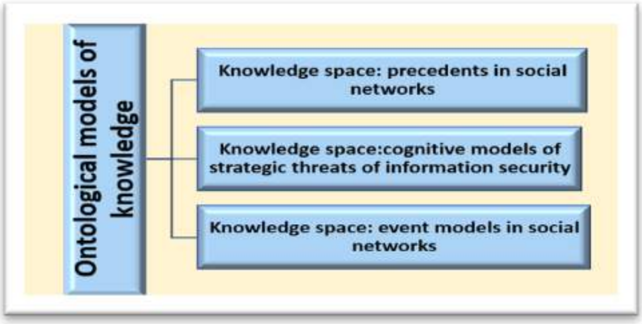

Ontological models of knowledge
Knowledge space: precedents in social networks
Knowledge space:cognitive models of strategic threats of information security
Knowledge space: event models in social networks

Figure 4. Ontological models of knowledge in the information security: infrastructure of social networks.

relationships, which generalizes the hierarchical data structure that is commonly used for filling in the task of extracting information and includes the steps (see Fig.4).

Indicators of the effectiveness of information extraction algorithms are divided into two classes, namely, the correctness indicators, for example, accuracy, correctness of the extracted information, completeness: the amount of information allocated relative to the volume of all available information and the measure of redundancy, as well as estimates of computing resources such as time and memory.

A query using ontologies can be performed automatically using the mechanisms of logical inference. For this we use the SPARQL language as the query language for ontologies.

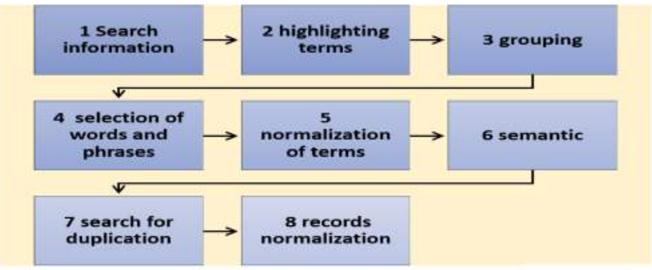


Figure 5. Ontological approach to the representation of knowledge: the allocation of information.

The choice of this language is due to the high level of its development, maturity and good potential, as confirmed by the following facts:

- SPARQL language received the status of official recommendation of the W3C2 consortium in 2008;
- SPARQL language is not tied to a specific program complex, unlike other query language ontologies;
- for SPARQL there is many software implementations and applications [4,5,12].

Below are examples of the use of the SPARQL language in scientific research:

Example 1. The list of topics of portal participants, which are actively explored within the framework of the subject matter of interest. Interpretation of the query: 'to issue all the results for the last (2016) year and sort them according to the descent of the occurrence in these results.' We formalize it in the language SPARQL. First, we will form a set of Terms containing all terms with / without repetitions), compared to the results of the activity for the year.

```
SELECT ?term
WHERE {
?term a cs:term.
?res a swrc:Result.
?res swrc:isAbout ?term . ?res swrc:year 2016. }
```

The resulting set of terms Terms, it is necessary to sort by descending the number of repetitions of each unique element. The terms at the beginning of the sorted list determine the directions that are actively exploring within the area of knowledge of interest.

Example 2.The list of users in the direction of interest. Interpretation of the query in SPARQL language is as follows: 'give a list of users whose search results are related to the terms of the given direction T t1,. . . , tn '. We formulate this query in SPARQL.

```
SELECT DISTINCT ?person
WHERE {
?person a swrc:Person.
?res a swrc:Result.
?res dc:creator ?person.
{?res swrc:isAbout t_1 }
UNION{?res swrc:isAbout t_2}...
UNION{?res swrc:isAbout t_n}.}
```

Example 3.List of publications similar to the given.Interpretation of the query in SPARQL as follows issue a list of the query associated with the terms that characterize the given search".The formal record of this query in SPARQL is presented below:

```
SELECT DISTINCT ?
WHERE {
?p a swrc:Publication . ?term a cs:term.
?p swrc:isAbout ?term . Pub swrc:is About ?term.}
```

Example 4. List of forums devoted to the direction of interest. Let's write this query as follows: 'give a list of forums related to the terms of the given direction $T\{t1, ..., tn\}$'. Interpreting the query in SPARQL:

```
SELECT DISTINCT ?forum
WHERE {
? forum a swrc: Forums.
{?forum swrc:isAbout t_1}
UNION{?forum swrc:isAbout t_2}...
UNION{?forum swrc:isAbout t_n}.}
```

The relationship between queries, the formal model of the system being developed, and the query code in SPARQL allows you to control the impact of:

- modifications of the set of requests received in the system and the ontologies used on the system code;
- modifications in the system code to the used ontologies and considered queries of the IS domain.

And creates additional opportunities for effective software verification at all stages of its life cycle [4,5,12].

IV. CONCLUSION

Social Networks on the Internet have become an important part of daily digital interactions for more than half billion users around the world. Unconstrained by physical spaces, the Social Networks offer to web users new interesting means to communicate, interact, and socialize. The Social Networks exhibit many of the characteristics of human societies in terms of forming relationships and how those relationships are used for personal information disclosure. However, current Social Networks lack an effective mechanism to represent social relationships of the users that leads to undesirable consequences of leakage of users' personal information to unintended audiences [20].As a result of the theoretical studies and their practical implementation, the ontology of the information security domain was developed in social networks, in particular, a method for identifying the user profile of the Internet was developed. Based on the research of the subject area, models and algorithms are built, architectural and technological decisions based on ontologies are developed to create a system of replenishment and storage, analysis and delivery, upon request, of information activity, information on

Web pages in social networks.The ontology is implemented together with the knowledge base - using ontologies and the SPARQL language, where a formal description of the queries to the system, creating guarantees for their computation and additional capabilities, provides effective verification of the system code at all stages of its life cycle. Such a structure is a distinguishing feature of the developed ontology and knowledge base. It allows you to efficiently process user requests.Be course of cybersecurity raises a host of questions about intellectual property protection and new tools and regulations have to be developed in order to solve this problem. A technological approach to protecting privacy might by cryptography although it might be claimed that cryptography presents a serious barrier to ICT criminal investigations. Therefore, it must be studied how people assign credibility to the information they collect to invent and develop new credibility systems to help consumers to manage the information overload.And the ontology for information security and the base of precedents were used to develop a software package designed to manage risks while ensuring network security, monitoring and preventing threats. We propose an ontological model to represent diverse social relationships and manage self-presentation of social web users.This model regulates personal information disclosure since social role and relationship quality between the users. We also present results of our user study, which demonstrates that relationship quality plays vital role to control personal information security disclosure in social networks on the Internet, and quality of relationship between users can be easily inferred from user interaction patterns in online social networks. The proposed method on the Internet user identification allows to achieve about 70 percent probability for systems monitoring the status of social networking resources.

References

[1] Vasenin V.A. To the creation of an international system for monitoring and analyzing the information space for the prevention and cessation of cyber conflicts // Information Technologies, 2012, No 9, 2-10 pp.

[2] Mirzagitov A.A., et al. Methods of developing an ontology for information security, based on the precedent approach, Vestn. Novosib. state. un-ta. Series: Information technology, 2013, vol.11, No 3, 37-46 pp.

[3] Domarev V.V. The security of information technology. Systems approach. K: DiaSoft, 2004, 992 pp.

[4] Aktayeva Al. & etc. - Technique of identification of users of social networks on ontologies // International scientific journal "Modern IT and IT - education", vol.12, No2, 2016, 26-34 pp.

[5] Aktayeva Al. & etc. - Cognitive Ontology of information security priorities in social networks // International Scientific and Practical Conference Open Semantic Technologies for Intelligent Systems - OSTIS-2017, 365-376 pp., http://proc.ostis.net/eng/ main.html

[6] Schumacher M. Security Engeneering with Patterns, LNSC 2754. Springer-Verlag Berlin Heidelberg, 2003, 87-96 pp.

[7] Jutla D.N., Bodorik P., Gao D. Management of Private Data: Web Services Addressing User Privacy and Economic, Social, and Ethical Concerns, in Secure Data Management. Toronto, Canada, 2004, 100-117 pp.

[8] Undercoffer J. Modeling Computer Attacks: An Ontology for Intrusion Detection. University of Maryland, Baltimore, 2004

[9] Stepanov P.A. Automation of the processing of natural language texts / / Vestnik NSU, Serie: Information Technology, 2013, vol. 11,No 2, 109-115 pp.

[10] Palchunov D.E., Stepanov PA Application of model-theoretic methods for extracting ontological knowledge in the domain of information security // Program Engineering, 2013, No 11

[11] Palchunov D. Ye., Yakhyayeva G. E., Hamutskaya A. A. The program system of information risk management RiskPanel // Program Engineering, 2011, No 7, 29-36 pp.

[12] Golomazov D.D. Selection of terms from the collection of texts with a specified thematic division / D.D. Golomazov // Information Technologies, 2010, No, 8-13 pp.

[13] Kolin K.K. Philosophy of information: the structure of reality and the phenomenon of information // Metaphysics, 2013, No, 61-84 pp.

[14] Ursul A.D. Nature of information: a philosophical essay. - 2 nd ed. - Chelyabinsk, 2010, 231 pp.

[15] ISO/IEC 27032: 2012. Security techniques. Guidelines for cybersecurity

[16] http://www.ontology-of-designing.ru/article /2014 4%2814%29 /7 Vorozhtsova.pdf

[17] https://core.ac.uk/download/pdf/53068632.pdf/data of the application 01.09.2017/

[18] https://online.zakon.kz/m/Document/?doc-id=39754354

[19] http://www.lib.tpu.ru/fulltext/v/Bulletin-TPU/2014/v324/i5/08.pdf/data of the application 01.09.2017/

[20] https://link.springer.com/chapter/10.1007/978-3-319-45880-9-5

[21] https://www.slideshare.net/guidelines-introduction-to-iso-27032

КОГНИТИВНЫЕ ОНТОЛОГИЧЕСКИЕ МОДЕЛИ КИБЕРБЕЗОПАСНОСТИ СОЦИАЛЬНЫХ СЕТЕЙ

А.Актаева
Кокшетауский университет имени А. Мырзахметова
г. Кокшетау, Казахстан
Р. Ниязова
Евразийский национальный университет им.
Л.Н.Гумилева
Астана, Казахстан
Е.Макатов
Кокшетауский университет имени А. Мырзахметова
г. Кокшетау, Казахстан
А.Даутов
Кокшетауский университет имени А. Мырзахметова
г. Кокшетау, Казахстан

Рассматривается структура и основные принципы технологии повышения вероятности идентификации субъектов информационных процессов открытых ресурсов сети Интернет на основе методов онтологии. На основе этой онтологии была реализована база знаний, предназначенная для создания программных систем, поддерживающих обеспечение информационной безопасности. Разработанная онтологическая база знаний была использована при разработке программного комплекса, предназначенного для идентификации пользователя социальных сетей при обеспечении информационной безопасности, отслеживания и предотвращения угроз.