BULGARIAN ACADEMY OF SCIENCES
PROBLEMS OF ENGINEERING CYBERNETICS AND ROBOTICS • 2021 • Vol. **75**, pp. 15-34
p-ISSN: 2738-7356; e-ISSN: 2738-7364
https://doi.org/10.7546/PECR.75.21.03

Problems of Human Secure Interaction with the Internet Space

Man Tianxing[1*], *Radoslav Yoshinov*[4], *Vasily Osipov*[2], *Nataly Zhukova*[1,2], *Mariya Schukina*[3], *Elena Evnevich*[2]

[1]*ITMO University, St. Petersburg, Russia;*
[2]*Laboratory of Computer-Information Systems and Software Engineering, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russia;*
[3]*St. Petersburg State Institute of Psychology and Social Work, St. Petersburg, Russia;*
[4]*Laboratory of Telematics, Bulgarian Academy of Sciences, Sofia, Bulgaria*
E-mail: mantx626@gmail.com

Abstract: The paper addresses the problem of ensuring secure human interaction with the Internet space by cognitive assistants. Psychological prerequisites to cognitive assistant creation are given consideration to, and the requirements the assistants should meet are formulated and substantiated. Feasibility studies on technological aspects of cognitive assistant creation are carried out. The methodology is proposed for the development of cognitive assistants with associative thinking. It involves the development of conceptual models of the human-Internet interaction process based on stream recurrent neural networks with controlled elements taking into account the psychological, physiological, and information aspects of the interaction. A series of experiments with recurrent neural networks was conducted.

Keywords: *Cognitive agents, HCI theory, Associative thinking, Assistive technologies, Cognitive assistants, Neural networks*.

1. Introduction

The Internet has become an essential factor in social life. 55.1% of the world's population has access to the network, which means more than 4.2 billion people (Miniwatts Marketing Group, 2019). The number of Internet users varies depending on the country. For example, the Internet access rate is estimated at

36.1%, 76.1%, 95% in Africa, in Russia, in North America, respectively. Nowadays, Internet users are the subjects of numerous neurocognitive psychological destructive influences. There is an urgent problem of ensuring the psychological security of the user preventing its behavior from victimization. The ability of users to identify, recognize, and to eliminate such threats by themselves is somewhat limited. And it is almost impossible to provide a personal psychologist to an individual to protect the psychics. All the above said affects not only the psychological and mental state of users but also the efficiency of interaction in the Internet space. In this regard, some authors propose to use a two-stage decision-making approach to survey the excessive usage of smart technologies (Borissova et al., 2020). Special attention is to be done on the access control that is closely connected with the security and protection of the information (Dikov, 2019). On the other hand, the most common vulnerabilities should be identified and to propose some prevention mechanism (Stankov & Tsochev, 2020). It is necessary to develop cognitive assistant machines with a wide range of cognitive functions which can operate like psychologist and at a reasonable cost. The approach to the creation of those cognitive machines according to the principles of bio similarity seems to be promising. The cognitive machine is to be developed as a kind of intelligent system "thinking" like a psychologist to ensure security for Internet users. One of the types of such machines could be associative-thinking cognitive assistants. To create a cognitive assistant machine with a wide range of capabilities for efficient human interaction with Internet space, a common methodology is required as well as specific models, methods, and technologies.

In this article, the psychological prerequisites and the methodology for creating cognitive assistants for secure human interaction with the Internet space are proposed. The main contributions of our work are summarized as follows:

1) This article discusses the premise of constructing a cognitive assistant from a psychological perspective. It addresses the problem of ensuring secure human interaction with the Internet space. According to the requirements of intelligent assistant creation, we carried out the feasibility studies on technological aspects of cognitive assistant creation.
2) The methodology for creating cognitive assistants with associative thinking is proposed. In the frame of the methodology, we designed the general functional scheme of assurance of psychologically secure human interaction with the Internet by cognitive assistants. The basic principles, methods, and technologies are formulated for the development and practical implementation of cognitive assistants. A structural core of the assistants based on an intelligent neural network is proposed.

2. Psychological prerequisites for creating cognitive assistants

The Internet has changed the way of a modern person's life, expanding the field of human cognition, creativity, communication, therapy, and other activities. The very nature of society has changed, which is now commonly called information society or digital one (although these concepts are not identical, and the first one is much wider than the second) (Emelin, 2018). The information generating, search and transfer are some of the main driving forces in modern society, and online tools play a significant role. The Internet has a wide range of impacts, a significant audience scale and a high rate of variability has become not only a space of interaction and self-expression but that of influence and control, the nature of which depends on information generating sources. Thus, the informatization of culture and public relations led to the emergence of new humanitarian and psychological risks. There are five types of the most widely spread ones:

1) Digital or cybercrimes. Among them: online personal data theft, financial resources theft, financial fraud, trafficking prohibited items for sale via Internet, human trafficking, involvement in destructive social groups, psychological impacts qualified as an inducement to commit suicide, sexual crimes.
2) Destructive communicative cyber-attack. It may destroy the user's social status, cause negative emotional experiences, self-attitude damage (reduced self-esteem, personality fragmentation, self-regulation violation, etc.), destruct important social connections, drive a person to suicide or homicide. An extensive typology of cyber-bullying is developed for a variety of manifestations of aggressive influences of a kind: cyber trolling, hating, flaming, cyber-stalkings, griefing, sexting (Bochaver & Khlomov, 2014). The most vulnerable to cyber-bullying is supposed to be the youth audience, which is distinguished for its openness to social contacts in the situation while the tools for recognizing an aggressive influence and for protecting against it are not formed yet. A recent survey of adolescents communication spheres revealed that strong social connections in online spaces are quite different from those in offline ones, namely: for the half of the adolescents their parents were excluded from the social circle in the network, and the presence of adult relatives is even more limited; at the same time, the strangers are easily accepted into virtual life of the adolescents, the latter not realizing that yesterday's strangers can be both sources of necessary connections and positive changes in the present and future and possible negative agents of influence as well (Soldatova & Chigarkova, 2018).
3) Privacy violations. Internet of Things, the collection and use of digital traces of users, the personal data and metadata dissemination through online

services and applications, the private information published by other users are examples of unintentional or deliberate violation of the right to privacy, of intrusion into personal space by users and generators of the Internet environment: nowadays privacy doesn't mean secrecy and anonymity, but the issues of privacy become rather associated with the problem of security and protection of personal data (Soldatova et al., 2017).

4) Personality disorders and developmental, mental deformities. For the world of childhood, these are the problems of transformation of the mental development of children who can't imagine life without gadgets and online tools. Among the problems there are changes in the character of the preschooler's leading type of activity – a role-playing plot game that is reduced under the influence of digital relations; there are more cases of lagging in the development of emotional and motivational needs, arbitrary regulation of behavior, lack of hand motor skills, impoverishment of the communicative sphere, in particular, emotional and non-verbal means of communication, etc., the appearance of the illusion of joint activity, the effect of the "telescope" (an exaggerated vision of the fragmentary identities of other people), the prevalence of clip presentation of information and clip thinking, the dominance of involuntary cognitive processes (attention, perception, memory, etc.) These effects of Internet influence lead to a weakening of regulatory processes (Moskvin & Moskvin, 2018), a general decrease in the arbitrariness of cognitive and behavioral acts. The distortion associated with the negative image of one's own "Self" and the reassessment of other images leads to such profound consequences as activation of dysfunctional attitudes in the form of an estimated dependence, obligation, dominance of comparative rather than genuine motivation (Zhuravleva, 2018). A person's life ceases to be for him "an effort in time" (M. Proust, 1927), independent work on himself, and the conscious formation of his life path. All this generates two global consequences: 1) the development of depressive-narcissistic experiences (emotional instability, unmotivated aggression, irritability, mood swings, apathy, atrophy of motivation); 2) problems of identity: misunderstanding, non-acceptance of oneself, loss of feeling of Self, loss of Self in time, the fragmented image of Self and painful experience of all these phenomena (Zhuravleva, 2018).
5) Digital addictions. Screen addiction, Internet addiction, dependence on social networks, online gaming addiction, online shopping addiction, online sexual addiction, etc. In general, the listed Internet addictions are understood as an obsessive (compulsive) desire to use the Internet or its

resources, leading to negative consequences in the professional, family, or bodily sphere of human life (Kolmogortseva, 2018).

In modern science, there is no common agreement on the independent status of the listed destructions. They are often supposed to be reflected or continued forms of prototypes or analogs from offline reality. However, under exposure to those risks, the vulnerability of Internet users is exacerbated by the fact that Internet reality is a relatively new space: objectively from the standpoint of historical experience and subjectively from the perspective of the life experience of a particular person. Behavioral and cognitive tools, emotional protection against Internet threats, and prevention of psychological security violations of users have not yet been sufficiently developed. Even children and adolescents who were born and were growing in the Internet era are not protected from those dangers because their mental and behavioral protection cannot be developed to a full extent due to age restrictions.

For the reasons mentioned above, it became more apparent than ever that development of a system of measures to ensure the security of Internet users in modern society is of urgent need. The response to the problem was a variety of legal, educational, psychological, and other solutions for creating a secure Internet environment (Baranov et al., 2017). From the part of information technologies, such an answer could be represented by new artificial intelligent systems – cognitive assistants. The following section describes the concept and creation methodology of such assistants.

3. Requirements for cognitive assistants

As an analysis of the known literature shows, requirements for such promising cognitive assistants, based on psychological aspects, have not been formulated previously. From the standpoint of modern psychology, cognitive assistant creation should proceed from two general principles:

- Cognitive assistants are designed to help users to recognize potential cyber-security threat while making decisions about operations with potentially dangerous content or information exposure; to recognize dangerous intentions of other users or dangerous content that those users want to pass to / to impose on the other ones;
- Cognitive assistants are designed to identify users who are more vulnerable to dangerous effects due to insufficient development of the hazard recognition skill, mental deficiency, or impairment due to age, illness, or personality traits (personal victimhood), resulting in poor hazard recognition and protection means choice.

From psychology, danger recognition in the Internet environment could be treated as a problem of how the user perceiving an object or a situation concludes

that the latter is dangerous. In psychology, there are several competing concepts of the organization of perception. Still, the leading role among them belongs to the so-called subject-oriented theories, where perception is understood as the process of making a decision (in our case by an Internet user) on assigning a certain category to an object (to classify an object as belonging to some category, to solve a categorization problem) (Gusev A.N., 2011). For the first time, this process was described in 1894 by the Russian psychologist N.N. Lange, who had found that with a consistent increase in exposure time a perceptual image went through several stages to a clear awareness of a particular image: first something was identified as being identical according to the weakest matching criteria, and then the image is refined step-by-step by strengthening those criteria (Allakhverdov, 2000).

In modern cognitive psychology, a whole group of concepts has been developed that considers perception as a solution to the categorization problem. American researcher of visual perception J. Gibson argued that a person perceives complex information but not individual stimuli. According to Gibson information is not what is fed to the receptor but what the observer obtains during perceptual activity, i.e., that aimed at the perception. The outstanding American cognitive psychologist J. Bruner treated perception as a categorization process – highlighting certain signs of a stimulus by which the perceived object is categorized. This process includes the elements of thinking since it implies putting forward and testing the hypotheses of objects belonging to some categories, which is a typical mental task. Perception is a step-by-step solution of the perceptual task, which includes primary categorization, search for signs of belonging to a category, confirming test of the belonging, and final categorization. In the presence of a set of alternative categorization hypotheses, the perceptual decision would correspond to the one with higher readiness, i. e. to the one which is more in line with the subject's expectations, for two main reasons: 1) the stimulus corresponds to the most habitual and frequent object appearances in the stimulus flow, that is, it is common; 2) the stimulus corresponds to the needs and emotional state of the object.

The author of the breakthrough work "Eye and Brain: The Psychology of Seeing" R. Gregory also emphasized that perception is the process of interpreting sensations, the process of creation, and testing hypotheses through preliminary experience. Perception is not only what we feel but also what we guess. The Nobel Prize winner psychologist D. Kahneman describes perception as a sequential decision-making process starting from sensory registration of a stimulus and finishing at its categorical interpretation which includes the following steps:

- Formation of primary perceptual units (preliminary selection of perceived objects in the field of attention).

- The figure accentuation in the attention field: choice of the figure and its units for analysis.
- Activation of identification units: attention and effort.
- Selection of interpretations (including those based on the principles of perceptual readiness).
- Answer selection / perceptual decision making.

Eminent cognitive scientist U. Neisser in his generalized theory of perception, introduces the concept of a cognitive scheme, which is a mental structure that anticipates acceptance of incoming information by the perceptual system. The scheme controls perceptual activities by choosing an object, by exploring its sensory qualities, by extracting information and thus resulting in modifying itself during this process. The scheme combines different levels of experience representation, including its details (the sensory basis of an image), specifics (the objective meaning of an image), emotional component (personal sense of an image).

Allakhverdov, Russian psychologist who proposed a generalized interpretation of the functioning of consciousness in the form of empirically proven laws being confirmed by numerous experimental research of more than a hundred years of the history of psychology, points out that in the process of choosing between alternative hypotheses several laws of consciousness are involved in the decision making (Allakhverdov, 2000):

- The Freud-Festinger rule shows that consciousness is trying not to notice, to ignore information that does not correspond to the usual experience or expectations in the situation (“it seemed to me”; “this cannot be because it can never be”); if it doesn't work and the stimulation continues, the consciousness offers to distort the interpretation of the situation in such a way as to bring it to the usual interpretation, to get rid of the contradiction with the usual one;
- The pattern interrupt explains that an unexpected change in context (versions, hypotheses on the perceived) causes emotional shock and failure in behavior until the work of consciousness aimed at situation reinterpretation results in the adoption of a new context or another hypothesis on the object under perception;
- The Hick's law: the more unexpected are the stimuli or reactions, the longer the consciousness works on them to move from one stereotype of interpretation (categorization) to another one.

As regards the task of recognizing potentially dangerous information in the Internet environment, the laws mentioned above make it possible to understand that in the information flow of the Internet environment, the user constantly solves the problem of recognizing stimuli in the system of categories “dangerous” and “secure”. Moreover, the solution is most likely to be represented in a continuous

scale, where a specific stimulus takes its place depending on the number of signs of its danger, rather than in a discrete system (dangerous or secure). For this mechanism to work efficiently, several conditions should be met:

- In the mind of the user a potential alternativeness to the interpretation of the content should be formed, that is, an understanding that it can be both secure and dangerous; this alternative is formed in the course of learning or gaining personal experience;
- In the mind of the user standards and patterns of the dangerous situation must be formed, that is, systems of signs, based on which the decision on belonging to the category of dangerous (hazardous) would be made; it is important to take into account the Wittgenstein-Roche law: class members are unequal, among them, there are more or less typical (Allakhverdov, 2000); hence the less the perceived object looks like a dangerous one, the more difficult it is recognizable as dangerous;
- To recognize dangerous content or impact as soon as possible the user must be ready to their presence in the Internet environment, that is, to be aware of their potential presence and to be in a state of anticipation of possible danger;
- To adjust to new challenges and risks of potentially dangerous content the reference categories (cognitive schemes, according to U. Neisser, or constructs, according to G. Kelly) should be penetrable, that is, maintain the potential of variability based on the user's readiness for an alternative reinterpretation of incoming information.

Due to the inability of Internet users themselves to solve the above-mentioned recognition, categorization, etc., tasks to a full extent, their solution should be delivered to cognitive assistants.

Another ability that a cognitive assistant should possess is identifying categories of Internet users being vulnerable to Internet-caused injuries. One of the most vulnerable categories is represented by Internet users with "the victim's complex" or the pronounced potential of a victim, referred to as victimhood in psychology. Victimhood is a complex of physical, mental, and social features and signs acquired by a person that can make him predisposed to become a victim of crime, accident, destructive cult, addiction, etc. (Malkina-Pykh, 2016). The first victimology studies in the mid of the twentieth century by G. Von Götting (Germany) and B. Mendelssohn (USA) showed evidence of the contribution of personal victimhood to the commission of crimes. Concerning the problem of cybersecurity, this means the importance of identifying potential victims and the priority of the introduction of cognitive assistants into the Internet activities of this risk group. And the different character of manifestation of the propensity to the role of the victim should be taken into account. Russian criminologist and victimologist D.V. Rivman identified several types of victims: universal one (high

potential vulnerability to any infringements), selective one (high potential vulnerability to certain types of crimes: financial, sexual, etc.), situational, casual and professional ones. Based on this typology, the zone of the introduction of cognitive assistants to prevent victimization of behavior should cover first of all universal and selective types of victims.

In modern psychology, there is a growing class of research aimed at identifying a potential victim position employing analyzing the digital traces of users and of identifying personal correlates with typical Internet behavior strategies. The pioneering work in this area by M. Kosinski and co-authors (Kosinski, Stillwell & Graepel, 2013) described predictive models enabling to determine sex, age, race, political preferences, religion, marital status, addiction to alcohol/drugs/smoking, etc. of the user with high probability by analyzing frequently visited Internet resources and affixed likes. Somewhat less strong but reliable were links with the parameters of intelligence, extraversion, openness to a new experience. About 70 Facebook likes contain sufficient information to predict user political preferences, attitude to health, and alcohol more accurately than his colleagues can do, and about 300 likes allow to do it better than the user's spouse could do (Youyou et al., 2015). They are tracking the position and movement of a person by GPS systems, semantic analysis of SMS, Internet requests and posts in networks – all that allows us to assess some psychological characteristics of a person more accurately than using standardized psychological questionnaires (Lambiotte & Kosinski, 2014). Besides, an increasing number of studies determine the personality traits and emotional state of users through content analysis of user texts (posts and comments), their images, and other content of personal pages (Bogolyubova, Upravitelev et al., 2018). Such an approach made it possible, for example, in a sample of Russian Facebook users to demonstrate a correlation between involvement in harmful online behavior and gender (higher in men) and psychopathic personality tendencies (Bogolyubova, Panicheva et al., 2018).

The considered concept of the creation of assistants with cognitive capabilities imposes quite strong requirements for possible technological solutions. The necessary capabilities are provided in the computer science domain.

4. Methods and technological possibilities of creating cognitive assistant

To explore the possibilities of realizing the formulated requirements for cognitive assistants to ensure safe interaction with the Internet space, we analyzed the well-known approaches to creating such systems.

The current intelligent cognitive assistants are often focusing on performing daily tasks, such as managing emails and answering common questions. The most

popular ones are Siri, Apple's virtual personal assistant; Mycroft, an open-source voice assistant, that can be installed on Linux, Raspberry Pi, or the Mark 1 hardware device; Cortana, a virtual assistant created by Microsoft for Windows 10, The users can interact with them through their embodied interface using natural language.

While the more significant focus is put on developing Intelligent Cognitive Assistants that perform everyday tasks, there also ones that target more domain-specific tasks: HealthPal (Komninos & Stamou, 2006), a personal medical assistant that helps to monitor various health conditions and alerts the user when there are abnormal indications. It also utilizes natural language to interact with the user; Cognitive Cockpit Assistant Systems (CASSY/CAMA) (Onken & Walsdorf, 2001) is developed and tested for aerospace and defence applications. It can increase the pilot's awareness of the situation in flight. The agent can understand the flight situation and goals, as well as the pilot's intentions and their possible errors. Based on these observations, the virtual assistant will interact with pilots in a human-like manner to help them understand the situation; (Coronado et al, 2018) proposed a modular cognitive agent architecture for question answering featuring social dialogue improved for a specific knowledge domain. The system has been implemented as a personal agent to assist students in learning the Java programming language; Daphne (Prat i Sala, 2017) is an intelligent cognitive assistant for reducing the cognitive load of the user by providing information, which is relevant to the current process of the search. This is done through the Critic Agent, which will offer criticism and feedback to the user about the design at hand.

A large amount of works is devoted to the development and application of artificial cognitive systems, including cognitive assistants, in the various domains of human activities (Oakley, 2018). As regards cognitive machine assistance, many theoretical and practical results are already obtained (Dormehl, 2017). However, those results are far from creating fully-fledged cognitive assistants in a broader sense. Mostly all cognitive abilities are limited to partial machine solutions of individual creative tasks for clearly defined situations. All known cognitive machine assistants are subdivided into three groups:

- Symbolic (those based on artificial intelligence)
- Emergent (those based on neural networks with associative signal processing)
- Hybrid ones.

For a long time, most attention was paid to symbolic cognitive artificial systems. It was possible to create a lot of practically operating cognitive machines. The common shortcomings of the systems of this group include the low level and complexity of their learning, poor creativity, the lack of associative information memorization, and others. These systems are used to solve clearly defined

individual creative tasks. None of the known systems of this type is cognitive in a broader sense. All attempts having been made for several decades to reach a higher level of cognitive abilities within the framework of a symbolic approach that has not been successful. Currently, the possibility of creating assistant machines being cognitive in a broader sense, i.e., capable of modeling a wide range of creative human functions, is associated with the development of bio-inspired emergent cognitive neural network machines. Examples of well-known emergent cognitive neural network solutions are the following: Hierarchical Temporal Memory (HTM); Deep SpatioTemporal Inference Network (DesTIN) system containing a hierarchical network of perception, similar to HTM, but more functional, and with a hierarchical network coordinated with it, responsible for actions and their reinforcement; Integrated Biologically based Cognitive Architecture (IBCA) is a large-scale emergent architecture that is intended for the modeling of distributed information processing in the human brain; Neurally Organized Mobile Adaptive Device (NOMAD) based on "neural Darwinism" theory; cognitive systems created on neuromorphic processors TrueNorth and others. However, no one of them has shown so far how to achieve high-level functions in cognitive neural network machines such as abstract thinking, complex conscious processing of video information, speech, and others. Without endowing cognitive assistants with artificial thinking, the realization of fully-fledged assurance of psychological security of Internet users with the help of those assistants seems to be problematic.

To solve the problem within the framework of the emergent approach, neural network structures are needed that are capable of fast and deep processing the perceived signals and of developing appropriate creative solutions. Recurrent neural network structures with associative information processing are likely to be the most suitable for. Nevertheless, until recent years the objective of meeting the requirements for both speed and depth of associative information processing hasn't yet been attained. In particular, Hopfield and Kosko's recurrent neural networks enable deep but not efficient information processing. Well-known real-time recurrent neural networks (RMLP, RTRN, Elman network) realize fast but not deep associative processing. All those neural networks in one way or another provide connection of the processed signals but are not able to realize the full-fledged associative thinking which is necessary for promising cognitive neural network assistants to possess. Some new solutions obtained in recent years (Osipov & Osipova, 2018) give hope for a breakthrough in this area in the course of their further development. To create associative thinking cognitive neural network assistants for the psychologically secure interaction of users with the Internet space it is necessary to enhance existing achievements, to search for fundamentally new neural network structures. At the same time the research is required on the issues of multilevel neural network coding and binding of processed signals in recurrent neural networks and on those of the development of methods for associative multilevel interactions control of heterogeneous

processed signals. Also new methods of learning of associative thinking cognitive assistants are required in order to solve problems of identifying, predicting and eliminating destructive neurocognitive influences on the users of the Internet space. Apart from the others there are issues of assessment of the functioning sustainability of such promising cognitive neural network assistants and of the creation of technologies for their software and hardware implementation. The following section proposes a new methodology for creating cognitive assistants.

By summarizing the possibilities and disadvantages of well-known approaches to creating intelligent cognitive assistants to ensure safe interaction with the Internet space, we can found that intelligent cognitive assistants have received significant development in terms of modeling the creative capabilities of a person and analysis of his condition in recent years. Implemented visual, mechanical, electrical, voice, and other types of interaction of these assistants with a person. The range of creative tasks assigned to cognitive assistants in various spheres of human activity has been substantially expanded, including some aspects of information and psychological security.

However, the ideas on which these assistants are built remain far from perfect. Basically, cognitive assistants function according to rules developed by humans. The capabilities of these assistants in perceiving, linking in space and time various events, and creating creative solutions are extremely limited. Great reliance on a breakthrough in this area is assigned to the development and application of self-learning multi-level neural network systems. So far, many of the neural network methods and technologies are reduced only to the recognition of static and dynamic images, filtering and predicting events, management based on recognized facts, but not too meaningful information processing and behavior.

To create promising cognitive assistants to ensure the safe interaction of a person with the Internet space, the capabilities of existing approaches are clearly not enough. A search for new solutions is needed to implement the advanced requirements for promising cognitive systems.

5. Methodology for creating cognitive assistants with associative thinking

In this section, we consider the proposed neural network approach to creating promising cognitive assistants with associative thinking, which allows them to be endowed with broad creative capabilities to ensure the safe interaction of a person with the Internet space.

To implement this methodology, we propose to use the intellectual neural network core, which has enhanced cognitive capabilities for processing heterogeneous information. Within the framework of the methodology under development the concept of the cognitive assistant with associative thinking implies the cognitive machine capable of perceiving events (signals), of their

space-temporal binding, of creating corresponding space-time models, of developing and realizing creative decisions (Osipov & Osipova, 2018). While creating cognitive assistants, it is planned to develop and to use conceptual models of human interaction with the Internet space. During model creation, the human-Internet interaction is investigated as the object of cognitive assistance from psychological, physiological, and information points of view. Revealing new behavioral patterns of the users and their typical psychological reactions to destructive influences can be provided by the series of experiments and subsequent processing of the obtained data by statistical and intelligent methods. Experiments include observation of processes of user-Internet interaction and of the changes in the user's psychological state depending on the perceived information.

The generalized functional scheme of the process of assurance of psychologically secure user-Internet interaction using the associative-thinking cognitive assistant is represented in Fig. 1.

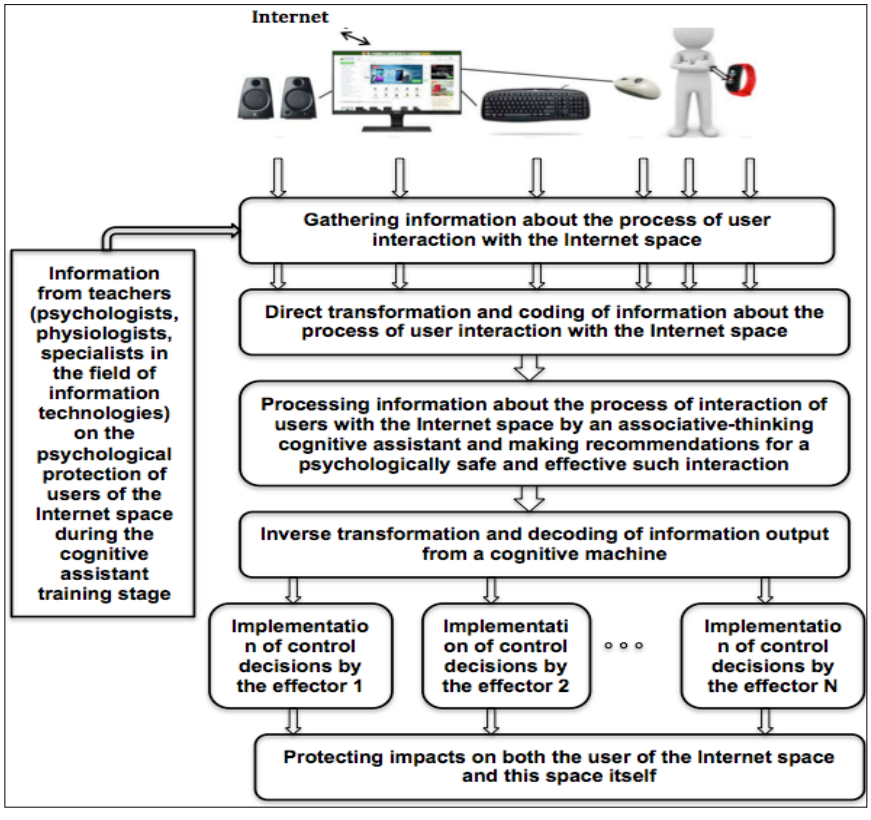

Fig.1. Generalized functional scheme of the process of providing the associative-thinking cognitive assistant for the psychologically secure interaction of users with the Internet space

The scheme proposed is based on the results having been obtained in the domains of psychology, physiology, intellectual information technologies. The development applies the existing methods of the analysis of a psycho-physiological state of a human, the general concepts of mentality, consciousness, thinking, methods of computer information processing. Except for the methods mentioned above the partial use is provided for the known methods of measurement of signals parameters, of coding, of binding and associative storing of information in neural network structures, for the methods of justification of expedient structures of intelligent systems, of the elements of the known cognitive architectures, for the methods of computer image recognition, forecasting and generation of the operating decisions. Heterogeneous information obtained from the Internet space and the user of this space (his expression, facial expressions, speech patterns, breathing, pulse, pressure, keyboard, and mouse actions) comes to the input of cognitive assistant.

For modeling the processes of ensuring the psychological security of users in their interaction with the Internet space, the system of the interconnected purposes and conditions of their achievement is created. Modeling is based on the new principles of machine intelligent associative assistance. The novelty of the principles consists of the fact that on the one hand, they enable taking into account new external factors, and on the other hand, they make it possible to identify and to apply new laws that characterize the interaction of users with the Internet space. The principles are based on some new patterns of interaction of elements in biological neural networks and on the new rules for multi-level binding of the processed signals. While modeling new mechanisms are used that provide multi-level space-time controlled associative processing of streams of heterogeneous signals that carry information about the process of user interaction with the Internet space. For the obtaining, pre-processing, and encoding information about the process of user interaction with the Internet space in the interests of cognitive assistance to the user, enhanced methods are used. Those methods allow us to recognize and to predict the state of mind of users of Internet space by taking into account their various emotions. Also, new methods are used for the learning of associative-thinking cognitive assistants so that they could detect destructive neurocognitive influences on the psychics of Internet users resulting in behavioral victimization. A separate group of new methods consists of methods of learning associative-thinking cognitive assistants to parry the destructive neuro-cognitive influences on the psychics of users of Internet space, protecting them from victimization.

For the application of the existing and new methods during the processing of the arriving information by the cognitive assistants, the effectors implement the

operating decisions. Effectors can make impacts both Internet users and on the Internet space itself.

In the processes of providing psychologically secure interrelations of users with the Internet space and of assessment of the adequacy of the associative-thinking cognitive assistants, several approaches are used: inductive and deductive ones and that of analogy with a human (biological) brain.

From a technical point of view, a cognitive assistant is an associative thinking machine, based on an intelligent neural network core. The proposed structure of the intelligent neural network core is shown in Fig. 2.

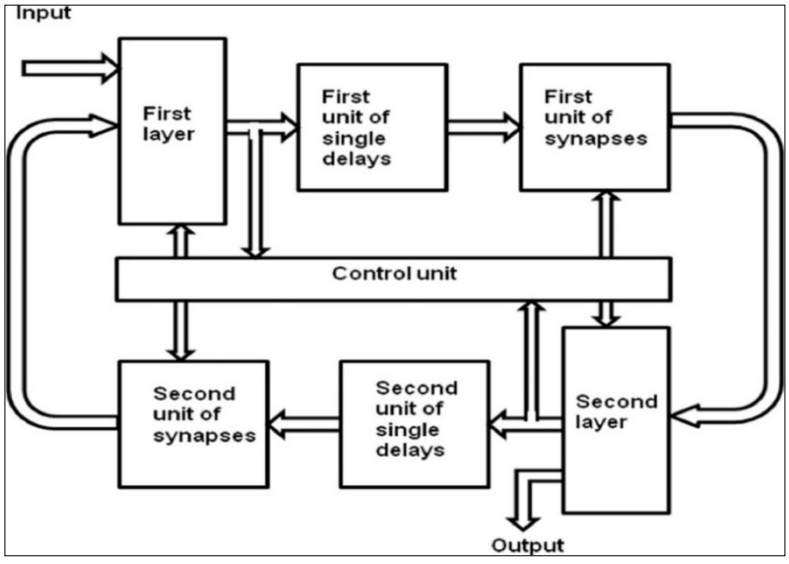

Fig.2 Structure of the intelligent neural network core

The neural network core is proposed to be developed by the advancement of the available results on the controlled associative and spatial processing of flows of heterogeneous signals in new artificial recurrent neural networks with controlled elements (flow recurrent neural networks). Those neural networks have rather enhanced capabilities for associative intelligent information processing. They can be endowed with various space-time structures and can provide controlled associative recurrent information processing.

The examples of such structures of recurrent neural network (RNN) are shown in Fig. 3 and Fig. 4.

Such networks and the associated methods of intelligent information processing can be significantly enhanced through the use of analogies with the human brain. The application of biological analogies involves the development of new methods of controlled space-time binding and intelligent processing of streams of heterogeneous signals in recurrent neural networks.

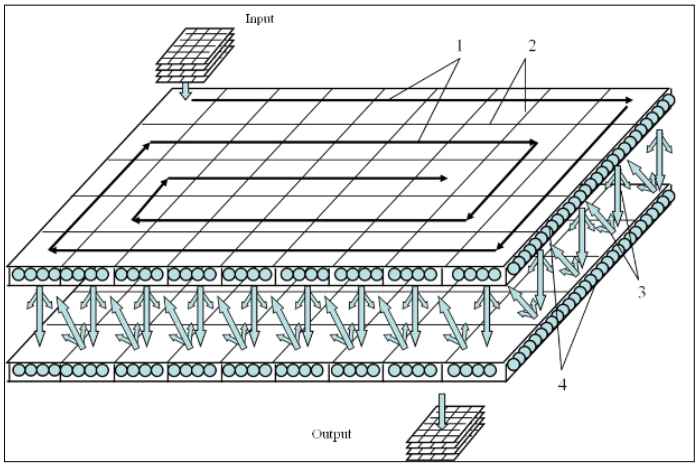

Fig. 3. The RNN structure in the form of a converging spiral: 1, 3 are direction of signals promotion along and between the layers of the RNN; 2 is the lines of splitting layers into logical fields due to spatial shifts of signals during transmission from one layer to another; 4 are neurons of the first and second layers

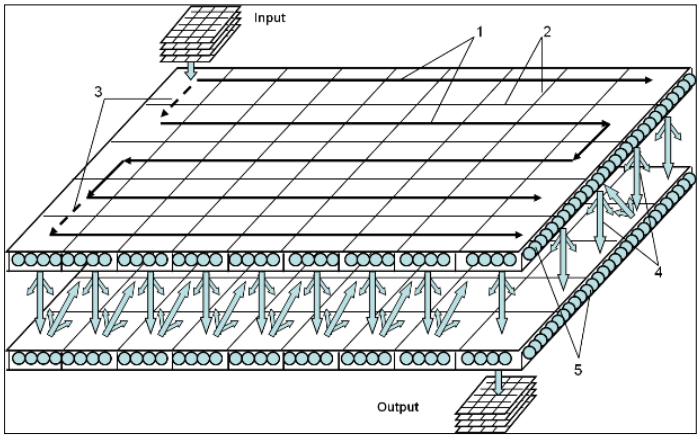

Fig. 4. An example of the three signaling systems RNN: 1, 3, 4 – directions of signals promotion along the layers and between them; 2 – lines, dividing the layers into logical fields due to spatial shifts of signals along the layers; 5 – neurons

The use of new neural networks allows us to provide associative-thinking cognitive neural network assistants with "inherited" information, providing them with enhanced cognitive capabilities as well as expanding the capabilities of the assistant memory and increasing the speed of parallel cognitive processing of incoming signals. Besides, through the creation of new neural networks assistants become capable of self-development and restoration of their associative structures.

New neural network solutions determine the need for the development of new architectures that are fundamentally different from the existing ones. The

search for architectures of associative-thinking cognitive neural network assistants with broader functions of the efficient assurance of psychologically secure human interaction with the Internet space is carried out following the formulated principles of machine associative assistance. The novelty of architecture consists of the flexible multilevel neural network organization providing the solution of various creative tasks connected with intelligent associative processing of heterogeneous information by means of the same neural network kernel. The development of architectures of associate thinking assistants is carried out based on a developed core by integrating several such cores. Due to kernel integration, it is possible to form complex, intelligent neural network structures based on a multitude of interconnected small neural network cores. New methods of identifying, predicting, and counteracting the destructive neuro-cognitive influences on the psychics and consciousness of users of the Internet space and the corresponding methods of accelerated neural network learning are being developed with an orientation to such architectures.

Particularly, in such an RNN with transparent internal logical structures, the tasks of direct and inverse signal analysis and synthesis problems can be solved. In the formation of multilevel space-time structures of the RNN, simultaneous associative signal processing is possible both in real time and in model time. The ability of a neural network machine to process signals both in real time and in model time allows us to talk about the presence of associative thinking.

Associative intelligent cognitive assistants with a new architecture can be developed in two versions: a personal one, hosted on the personal computer itself, and group, which is accessed via the network.

The fundamental differences between the used intellectual neural network core and the known solutions are the following advanced capabilities of associative information processing:

- Recognize and predict events with lifelong learning, also refer to the past.
- Synthesize processes according to their known initial and final events.
- Change the sequence of events and processing time, switch from one creative task to another.
- Solve various creative tasks on the same intellectual neural network core without restructuring its structure.
- Purposefully call from the associative memory the most related events with input signals, etc.

This core can be endowed with various internal logical structures that allow for more complete spatiotemporal binding of processed signals. Within its framework, several signaling systems can be implemented that are responsible for their levels of controlled associative processing. By controlling the associative processing of signals in such a recurrent neural network, depending on the current

state of the layers, various cognitive functions are realized: recognition, prediction, the reflection of the past, generation of “thoughts” to eliminate emerging information contradictions, and other functions.

4. Conclusion

The article formulates problematic questions on the development of such intelligent neural network cores in relation to ensuring the psychological safety of human interaction with the Internet space. The value and the novelty of the conducted research consist of the new ideas and approaches which enable to solve the problems of assurance of secure human- Internet interaction at the machine level. The development of such ideas required deep study of the current state and of the needs and capabilities in the various areas of knowledge, first of all, in the area of psychology. The situation analysis resulted in the understanding of the necessity of the creation of the fundamentally new intelligent systems – biosimilar, associative thinking cognitive assistants. A promising approach to their creation is based on the artificial recurrent neural networks with controlled elements. The methodology is proposed for the creation of cognitive assistants. In the frames of the methodology, both generalized functional scheme of assurance of psychologically secure interaction with the Internet by cognitive assistants is determined, and the structure of an intelligent neural network core of assistants is developed. The basic principles, methods, and technologies are formulated for the development and practical implementation of cognitive assistants. The experiments with RNN within the framework of the proposed approach confirm its validity.

In general, the conducted research gives a chance to the creation of the assistants capable of providing secure interaction of users and the Internet space. Broader application of assistants would result in qualitative changes of a state of Internet space and of the users' role in it.

It is crucial however, to pay attention to some risks of introducing digital cognitive assistants. Those are the risks associated with the personal features of developers and distributors of artificial assistants. Who, for what purposes, and on what grounds will determine the dangerous and secure zones? Who will make decisions on their boundaries and criteria? Are they identical or different for different categories of users? Here the contribution of voluntarism is rather large, and the problem of a measure of the subjectivity of different interaction participants categories is rather critical. Also, cognitive assistants, along with other digital tools, can turn into a kind of mental “prostheses” and “digital surrogates” of consciousness, causing a person to outsource personal tasks and personal choices to “smart” machines (Emelin, 2018). To avoid this undesirable development trends of artificial assistants it is important that they would become to a greater extent the tools of personal training as a full-right subject of Internet

interaction with responsible self-presentation, self-determination and independent decisions making and wouldn't be the tool replacing or replenishing the mental resources of a human.

References

1. Allakhverdov, V.M.: Сознаниекакпарадокс [Consciousness is like a paradox]. СПб.: Издательство ДНК, 528 с. (2000).
2. Baranov, N.A., Vossoevich, A.L., Vinogradova, S.M., Kefeli, I.F., Osipov, V.Yu.: Информационно-психологическая и когнитивная безопасность [Information-psychological and cognitive security]. Коллективная монография / Ed. I.F. Kefeli, R.M. Yusupov. ИД "Петрополис", СПб. 300 (2017).
3. Bochaver, A., Khlomov, K.: Cyberbullying: harassment in the space of modern technologies. Psychology. Journal of Higher School of Economics, 11(3), 178–191 (2014).
4. Bogolyubova, O., Panicheva, P., Tikhonov, R., Ivanov, V., Ledovaya, Y.: Dark personalities on Facebook: Harmful online behaviors and language. Computers in Human Behavior, 78, 151–159 (2018).
5. Bogolyubova, O., Upravitelev, P., Churilova, A., Ledovaya, Y.: Expression of Psychological Distress on Instagram Using Hashtags in Russian and English: A Comparative Analysis. SAGE Open, 8(4), 2158244018811409 (2018).
6. Borissova, D., Dimitrova. Z., Garvanova. M., Garvanov. I., Cvetkova. P., Dimitrov. V., Pandulis. A.: Two-stage decision-making approach to survey the excessive usage of smart technologies. Problems of Engineering Cybernetics and Robotics 73, 3–16 (2020).
7. Coronado, M., Iglesias, C. A., Carrera, A., Mardomingo, A.: A cognitive assistant for learning java featuring social dialogue. International Journal of Human-Computer Studies, 117, 55–67 (2018).
8. Dikov, D.: Telecommunications in Management Systems. Chapter 4 Security and Protection of Information in Management Systems, 185–309 (2019).
9. Dormehl, L.: Thinking Machines: The Quest for Artificial Intelligence--And Where It's Taking Us Next. Tarcherperigee (2017).
10. Emelin, V.A.: Технологический радикализм: обратная сторона «цифрового» Я [Technological radicalism: the reverse side of the "digital" I]. Личность в эпоху перемен: mobilis in mobili: Материалы международной научно-практической конференции 2018 / Ed. E.Yu. Patyaeva, E.I. Shlyaginoy. M., 112–114 (2018).
11. Gusev, A.N.: Sensation and perception.Russia: Центр "Академия" Press (2011).
12. Kolmogortseva, A.A.: Интернет-аддикция: зависимость от интернета или зависимость в среде нитернет? [Internet addiction: dependence on the Internet or dependence in the medium of Internet?]. Личность в эпоху перемен: mobilisinmobili: Материалы международной научно-практической конференции 2018 / Ed. E.Yu. Patyaeva, E.I. Shlyaginoy. M.: Смысл, С, 122–125 (2018).
13. Komninos, A., Stamou, S.: HealthPal: An intelligent personal medical assistant for supporting the self-monitoring of healthcare in the ageing society. In Proceedings of UbiHealth (2006).

14. Kosinski, M., Stillwell, D., Graepel, T.: Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences, 110(15), 5802–5805 (2013).
15. Lambiotte, R., Kosinski, M.: Tracking the digital footprints of personality. Proceedings of the IEEE, 102(12), 1934–1939 (2014).
16. Malkina-PykhI. G.: Виктимология. Психологияповеденияжертвы [Victimology. Psychology of victim behavior]. (Учебное пособие) - СПб.: Питер, 832 с. (2016).
17. Miniwatts Marketing Group, [Online] URL: https://miniwatts.com/, 2021/04/01.
18. Moskvin, V.A., Moskvin, N.V.: Индивидуальные особенности личности при проявлениях интернет-аддикции [Individual personality traits in manifestations of Internet addiction]. Личность в эпоху перемен: mobilis in mobili: Материалы международной научно-практической конференции 2018 / Ed. E.Yu. Patyaeva, E.I. Shlyaginoy. M.: Смысл, С., 125–127 (2018).
19. Oakley, D.: Brain and Mind - Animal awareness, consciousness and self-image. 1st edn. Routledge, UK (2018).
20. Onken, R., Walsdorf, A.: Assistant systems for aircraft guidance: cognitive man-machine cooperation. Aerospace Science and Technology, 5(8), 511–520 (2001).
21. Osipov, V., Osipova, M.: Space-time signal binding in recurrent neural networks with controlled elements. Neurocomputing, 308, 194–204 (2018).
22. Prat i Sala, A.: Daphne: An Intelligent Cognitive Assistant for the design of Earth observing satellites. Master's thesis, Universitat Politècnica de Catalunya (2017).
23. Proust, M.: À la recherche du temps perdu [In Search of Lost Time]. Grasset, (1927).
24. Soldatova, G.U., Priezheva A.A., Olkina O.I., Hatter V.N.: Practical psychology of security: personal data management on the Internet: a teaching aid for employees of the general education system. Russia: Federal Institute for Educational Development Press (2017).
25. Soldatova, G.U., Chigarkova, S.V.: Сильные и слабые связи в структуре социального онлайн-капитала у разных поколений [Strong and weak ties in the structure of social online capital among different generations]. Личность в эпоху перемен: mobilis in mobili: Материалы международной научно-практической конференции 17-18 декабря 2018 / Ed. E.Yu. Patyaeva, E.I. Shlyagin (2018).
26. Stankov, I., Tsochev, G.: Vulnerability and Protection of Business Management Systems: Threats and Challenges. Problems of Engineering Cybernetics and Robotics, 72, 29–40 (2020).
27. Youyou, W., Kosinski, M., Stillwell, D.: Computer-based personality judgments are more accurate than those made by humans. Proceedings of the National Academy of Sciences, 112(4), 1036–1040 (2015).
28. Zhuravleva, Yu.V.: Феномен современной неклинической деперсонализации: кризис идентичности в условиях медиапространства [The phenomenon of modern non-clinical depersonalization: a crisis of identity in a media environment]. Личность в эпоху перемен: mobilis in mobili: Материалы международной научно-практической конференции 2018 / Ed. E.Yu. Patyaeva, E.I. Shlyaginoy. M.: Смысл, С., 117–120 (2018).