NATIONAL ACADEMY OF SCIENCES OF UKRAINE
INTERNATIONAL RESEARCH AND TRAINING CENTER FOR INFORMATION TECHNOLOGIES AND SYSTEMS OF THE NAS OF UKRAINE AND OF MES OF UKRAINE
V.M. GLUSHKOV INSTITUTE OF CYBERNETICS OF THE NAS OF UKRAINE

Cybernetics and Computer Engineering

2 (208)/2022

SCIENTIFIC JOURNAL ▪ FOUNDED IN 1965 ▪ PUBLISHED 4 TIMES PER YEAR ▪ KYIV

CONTENTS

NATIONAL ACADEMY OF SCIENCES OF UKRAINE
INTERNATIONAL RESEARCH AND TRAINING CENTER FOR INFORMATION TECHNOLOGIES AND SYSTEMS OF THE NAS OF UKRAINE AND OF MES OF UKRAINE
V.M. GLUSHKOV INSTITUTE OF CYBERNETICS OF THE NAS OF UKRAINE

Cybernetics and Computer Engineering

2 (208)/2022

SCIENTIFIC JOURNAL ▪ FOUNDED IN 1965 ▪ PUBLISHED 4 TIMES PER YEAR ▪ KYIV

CONTENTS

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
МІЖНАРОДНИЙ НАУКОВО-НАВЧАЛЬНИЙ ЦЕНТР
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА СИСТЕМ НАН УКРАЇНИ І МОН УКРАЇНИ
ІНСТИТУТ КІБЕРНЕТИКИ ІМ. В.М. ГЛУШКОВА НАН УКРАЇНИ

Кібернетика та обчислювальна техніка

2 (208)/2022

НАУКОВИЙ ЖУРНАЛ ▪ ЗАСНОВАНИЙ У 1965 р. ▪ ВИХОДИТЬ 4 РАЗИ НА РІК ▪ КИЇВ

ЗМІСТ

DOI: https://doi.org/10.15407/kvt208.02.005

RACHKOVSKIJ D.A.[1, 2], DSc (Engineering),
Chief Researcher, Dept. of Neural Information Processing Technologies,
Visiting Professor, Dept. of Computer Science, Electrical and Space Engineering,
https://orcid.org/0000-0002-3414-5334, e-mail: dar@infrm.kiev.ua
GRITSENKO V.I.[1], Corresponding Member of NAS of Ukraine,
Directorate Adviser, ORCID ID 0000-0002-6250-3987, e-mail: vig@irtc.org.ua
VOLKOV O.Ye.[1], PhD (Engineering), Director,
https://orcid.org/0000-0002-5418-6723, e-mail: alexvolk@ukr.net
GOLTSEV A.D.[1], PhD (Engineering), Senior Researcher,
Acting Head of the Dept. of Neural Information Processing Technologies,
https://orcid.org/0000-0002-2961-0908, e-mail: root@adg.kiev.ua
REVUNOVA E.G.[1], DSc (Engineering),
Senior Researcher, Dept. of Neural Information Processing Technologies,
https://orcid.org/0000-0002-3053-7090, e-mail: egrevunova@gmail.com
KLEYKO D.[3], PhD (Computer Science), Researcher,
https://orcid.org/0000-0002-6032-6155, e-mail: denis.kleyko@ri.se
LUKOVICH V.V.[1]
Researcher of the Dept. of Neural Information Processing Technologies
ORCID ID 0000-0002-3848-4712, e-mail: vvl97@ukr.net
OSIPOV E.[2], PhD (Computer Science), Professor,
https://orcid.org/0000-0003-0069-640X, e-mail: evgeny.osipov@ltu.se
[1] International Research and Training Center for Information
Technologies and Systems of the NAS of Ukraine and of the MES of Ukraine,
40, Acad. Glushkova av., Kyiv, 03680, Ukraine
[2] Department of Computer Science, Electrical and Space Engineering,
Lulea University of Technology, 971 87 Lulea, Sweden
[3] RISE Research Institutes of Sweden AB,
164 40 Kista, Sweden

NEURAL DISTRIBUTED REPRESENTATIONS FOR ARTIFICIAL INTELLIGENCE AND MODELING OF THINKING

Introduction. *Current progress in the field of specialized Artificial Intelligence is associated with the use of Deep Neural Networks. However, they have a number of disadvantages: the need for huge data sets for learning, the complexity of learning procedures, excessive specialization for the training set, instability to adversarial attacks, lack of integration with knowledge of the world, problems of operating with structures known as binding or composition problem. Over-*

coming these shortcomings is a necessary condition for advancing from specialized Artificial Intelligence to general one, which requires the development of alternative approaches.

The purpose of the paper *is to present an overview of research in this direction, which has been carried out at the International Center for 25 years. The approach being developed stems from the ideas of N. M. Amosov and his scientific school. Connections to the Hyperdimensional Computing (HDC) and Vector Symbolic Architectures (VSA) field as well as to current brain research are also provided.*

Results. *The concept of distributed data representation is outlined, including HDC/VSA that are capable of representing various data structures. The developed paradigm of Associative-Projective Neural Networks is considered: codevector representation of data, superposition and binding operations, general architecture, transformation of data of various types into codevectors, methods for solving problems and applications.*

Conclusion. *An adequate representation of data is one of the key issues within the Artificial Intelligence. The main area of research reviewed in this article is the problem of representing heterogeneous data in Artificial Intelligence systems in a unified format based on modeling the neural organization of the brain and the mechanisms of thinking. The approach under development is based on the hypothesis of distributed representation of information in the brain and allows representing various types of data, from numeric values to graphs, as vectors of large but fixed dimensionality.*

The most important advantages of the developed approach are the possibility of natural integration and efficient processing of various types of data and knowledge, a high degree of parallel computing, reliability and resistance to noise, the possibility of hardware implementation with high performance and energy efficiency, data processing based on associative similarity search — similar to how human memory works. This allows one to unify the methods, algorithms, and software and hardware for Artificial Intelligence systems, increase their scalability in terms of speed and memory with an increase in data volume and complexity.

The research creates the basis for overcoming the shortcomings of current approaches to the specialized Artificial Intelligence based on Deep Neural Networks and paves the way for the creation of Artificial General Intelligence.

Keywords: *distributed data representation, associative-projective neural networks, codevectors, hyperdimensional computing, vector symbolic architectures, artificial intelligence.*

INTRODUCTION

The current progress in the field of specialized Artificial Intelligence is associated with the use of Deep Neural Networks. However, they have a number of disadvantages: the need for huge data sets for learning, the complexity of learning procedures, excessive specialization for the training set, instability to adversarial attacks, lack of integration with knowledge of the world, problems of operating with structures known as binding or composition problem. Overcoming these shortcomings is a necessary condition for advancing from the specialized Artificial Intelligence to the general one, requiring the development of alternative approaches.

In 1960s, Nikolai M. Amosov formulated a hypothesis [1] about the mechanisms of information processing by the human brain that produce intelligent behavior. Those ideas were further developed in his subsequent works, including [2–5]. In fact, an approach was proposed to create Artificial Intelligence based on modeling the principles of human thinking and neural network organization of the brain. To develop and implement the approach, at the turn of the 1960s, the Department of Biological Cybernetics was founded at the Glushkov Institute of Cybernetics. Since 1997, the work of Amosov's school has continued at the department of Neural Information Processing Technologies of the International Research and Training Center for Information Technologies and Systems of the National Academy of Sciences and the Ministry of Education and Science of Ukraine (the International Center).

Fig. 1. The B-512 neurocomputer that had a 512-bit machine word length.

Initially, the developments of localist semantic M-networks [2–5] and assembly Hebb-like neural networks were carried out in parallel at the Amosov's department, resulting in the world's first autonomous robot controlled by neural networks in a natural environment [3]. In the late 1980s, Ernst M. Kussul proposed the foundations of the original paradigm of the Associative-Projective Neural Networks (APNNs) [5, 6]. The idea was to combine the hierarchical organization of Amosov's world model with the advantages of distributed representations, as well as with Hebb's cell assemblies. For the efficient implementation of APNNs, as a result of two projects in Japan jointly with Wacom, high-performance specialized neurocomputers were created using the Japanese element base, see Fig. 1. This development entered the history of Ukrainian informatics.

The article presents an overview of the research that have been carried out at the International Center for 25 years in the direction of developing the ideas of N.M. Amosov and his scientific school. Therefore, it is important to emphasize that this article is focused heavily on the results obtained from a single research group and, hence, it does not give the full credit to the related ideas and methods developed by other groups. We highlight some connections to Hyperdimensional Computing (HDC) and Vector Symbolic Architectures (VSA), as well as to brain research, however, for a comprehensive treatment of HDC/VSA and its connection to APNNs the readers are kindly referred to [7, 8].

DISTRIBUTED DATA REPRESENTATIONS

To represent data of various type, modality, and complexity, APNNs use distributed representations. They are based on modeling a “distributed” or “holographic” representation of information in the brain, as an alternative to “localist” representations [9]. In localist representations, each “object” (for example, a feature, a physical object, a relation, a complex structure, etc.) corresponds to a certain node (neuron), represented by a vector component, the dimension of which is equal to the number of neurons. In the distributed approach, an object is represented as “distributed” over a set of neurons. Distributed representation is a form of vector representation where each object is represented by a subset of vector components, and each vector component can belong to representations of many objects. In distributed representation, the state of individual components of representation cannot be interpreted without knowing the states of other representation components. To be useful in applications, the distributed representations of similar objects must be similar (by some measure of similarity of the corresponding vectors — for example, by the value of the dot product or cosine of the angle between the vectors).

Distributed representations possess the following advantages:

– high information capacity. For example, if one object is represented by M binary components of a D-dimensional vector, then the number of representable objects is equal to the number of combinations M from D, in contrast to D/M in localist representations;

– direct access to the representation of the object. A distributed representation of a complex structure can be processed directly, without tracing pointers or connections required in symbolic or localist representations;

– an explicit representation of similarity. Similar objects have similar representations that can be directly compared using efficiently computable vector similarity measures (e.g., dot product, Minkowski distance, etc.);

– a rich semantic basis, which is provided through the direct use of representations based on features and the possibility of representing the similarity of the features themselves in their vector representations;

– for many types of distributed representations – the ability to recover the original representations of objects;

– the ability to work in conditions of noise, malfunction, and uncertainty, as well as neurobiological plausibility.

Since distributed representations of various objects are vectors, a rich arsenal of methods developed for vector data can be applied to their processing.

It was believed that the main disadvantage of distributed representations is the inability to represent the structure [10]. However, various researchers have developed a number of "structure-sensitive" distributed representations in various formats [10–13]. Such distributed representations are called hypervectors, and models based on them are called Hyperdimensional Computing (HDC) or Vector Symbolic Architectures (VSA) [7, 8]. The dimension of hypervectors is large, usually more than 1000, and reaches hundreds of thousands or more. The main operations on hypervectors are superposition, used to combine multiple hypervectors, and binding, used to associate them in representing structures. In various hypervector models, these operations are implemented in different ways, but the dimensionality of hypervectors at the input and output of these operations does not change.

THE ASSOCIATIVE-PROJECTIVE NEURAL NETWORKS

In APNNs, we use binary hypervectors with components from $\{0,1\}$. Moreover, those are sparse vectors, that is, the proportion of their non-zero components is small. This data representation format allows an efficient processing. Historically, and to distinguish from other hypervector types, we call such hypervectors as "codevectors", and will use the term throughout the text of this article.

As a superposition operation, component-wise disjunction of codevectors is used. When a set of codevectors is represented by a component-wise disjunction, the presence of some single component in the resulting codevector is determined by the

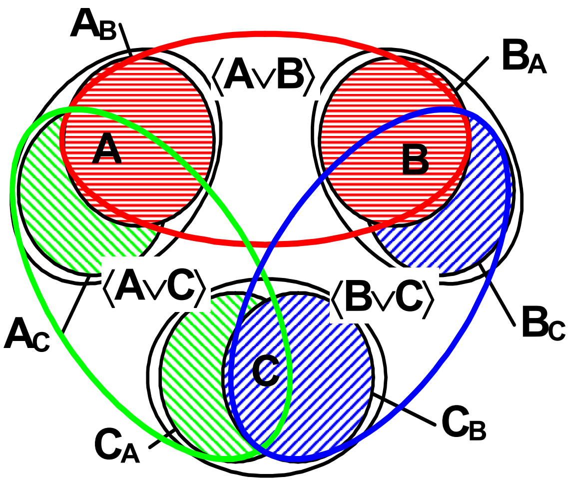


Fig. 2. Binding via the Context-Dependent Thinning. Smaller ovals represent the 1-components of the codevectors **A**, **B**, **C**. Larger ovals show the codevectors ⟨**A** ∨ **B**⟩, ⟨**B** ∨ **C**⟩, ⟨**A** ∨ **C**⟩ bound by CDT. Circles $\mathbf{X}_\mathbf{Y}$ denote the subset of 1s preserved in the codevector **X** when **Y** is also present. It can be seen that, e.g., $\mathbf{A}_\mathbf{B}$ and $\mathbf{A}_\mathbf{C}$ are different subsets of **A.** Note that actually the 1-components belong to randomly generated codevectors.

presence of such a component in at least one codevector of the set. Thus, the individual components do not contain information about the combination of codevectors in the set. Binding operations are used in HDC/VSA to preserve this information.

For binding codevectors, Context-Dependent Thinning procedures were proposed [10]. In one version of this procedure, the bound codevector ⟨**Z**⟩ is formed from the codevectors $\mathbf{X}_i$ to be bound as follows:

$$\mathbf{Z} = \vee_i \mathbf{X}_i;\ \langle \mathbf{Z} \rangle = \vee_{k=1,K} (\ \mathbf{Z} \wedge \mathbf{Z}^{\sim}(k)\ ) = \mathbf{Z} \wedge \vee_{k=1,K} \mathbf{Z}^{\sim}(k).$$

Here $\mathbf{Z}^{\sim}(k)$ is **Z** with permuted components. For each k, a random independent permutation is used which is fixed for the specific k. It is also possible to use the same permutation multiple times.

The subset of 1-components of each codevector $\mathbf{X}_i$ that is preserved in ⟨**Z**⟩ depends on **Z**, and hence on each and all $\mathbf{X}_i$, as shown in Fig. 2. Thus, information is stored about a specific set of elements of the set, the codevectors of which participated in the formation of ⟨**Z**⟩, ensuring binding.

The number K of used permutations controls the number of 1-components in the final ⟨Z⟩. Thus, it is possible to normalize the number of 1s, i.e., degree of sparsity, in the resulting codevector. Note that the operation of component-wise conjunction also provides binding, but increases the degree of sparseness of the resulting codevector, that might be an undesirable feature of binding operation.

The general APNN architecture. The APNN architecture was proposed and developed in [6, 14–18]. It is generally recognized that for a reasonable common-sense behavior, an intelligent agent needs a model of the world that includes knowledge specific to the subject area, as well as information about the agent itself. Such a model allows the agent to understand the world and helps it in its interaction with the world, for example, by predicting the results of actions. The goal of developing APNNs is to offer an approach to create a complex hierarchical model of the world that supposedly exists in the brain of humans and higher animals, as a step towards the Artificial General Intelligence. Two types of hierarchy are considered in APNNs: compositional (part-whole), as well as classification or generalization hierarchy (class-member or is-a). An example of a compositional hierarchy: letters → words → sentences → paragraphs → text. An example of a classification hierarchy: apple → big red apple → this big red apple in the hand.

The APNN model of the world is based on models of objects of various modalities, including sensory (visual, acoustic, tactile, motor, etc.) and more abstract modalities (linguistics, planning, reasoning, abstract thinking, etc.), which are hierarchically organized. Models should exist for objects of different nature, for example, events, objects, feelings, features, etc. Models (their representations) of different modalities can be combined, which leads to multimodal representations of objects and their associations with behavioral schemes (reactions to objects or situations), see [6, 14–18].

The APNN architecture is based on fixed-dimensional codevectors for objects of varying complexity and generality, which represent various heterogeneous data types, for example, numeric data, images, sequences, graphs (Sec. 4). Codevectors can be formed "on the fly" (without training). APNNs have a multi-module, multi-level and multi-modal structure, see Fig. 3 and [18] for more details. The modules are connected by the bundles of projective one-to-one connections that just copy codevectors between the modules. The architecture includes the modules of independent modalities and sub-modalities. For example, independent visual features of an object such as shape, size, texture, color are represented by codevectors in the modules M11...M14. A codevector of a visual model of an object is formed from these feature codevectors in the module M21. In the modules M22...M25, the codevectors of acoustical, tactile, olfactory, taste models are formed. The codevector of an integral multi-modality sensory model of an object is formed in the module M31, to which its name codevector could be added from the module M32. Probably, name could be a feature in the models of all modules.

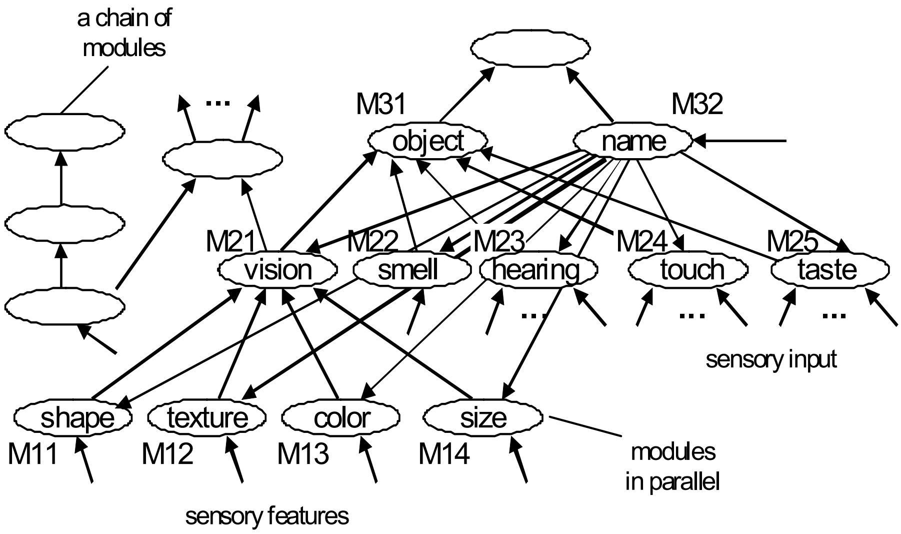


Fig. 3. An example of the APNN architecture.

The modules form, store and process a set of codevectors representing object models of a certain modality and a certain level of the compositional hierarchy. Module's codevectors are constructed from codevectors obtained from other modules, such as lower-level modules of the same modality, or from modules of other modalities. The lower level of the compositional hierarchy consists of modules that provide interface (of representations) with the external environment. A codevector is similar to the codevectors of its elements of the lower levels of the compositional hierarchy, as well as to the codevectors of higher levels, for which the codevector is an element. Thus, using similarity search in the memory of lower- and higher-level elements, it is possible to recover both the codevectors of the elements of lower levels and the compositional codevectors of higher levels. Similarity search in the memory of single module allows finding similar objects.

So, each module should have a long-term memory where it stores its codevectors. It was proposed to use distributed auto-associative memory of the Hopfield type as the module's memory [19, 20, 21, 22]. Let us consider it in some more details.

Associative memory and the generalization hierarchy. One of the key modes of processing codevectors is similarity search, i.e., search in the database (set) for a codevector most similar to the query codevector. This can be effectively done using a neural network distributed auto-associative memory with binary connections (an auto-associative version of the Willshaw memory). Each memory neuron (corresponding to the codevector component) is connected to all other neurons by a connection with a weight of 0 or 1. Each codevector is memorized according to the binary version of the Hebbian rule: the weight of the connection between memory neurons corresponding to the 1-components of the codevector is set to 1. If the weight of the connection is already 1, it does not change.

After the set of codevectors is memorized, a codevector-query is given as the input, which may not belong to the stored set of codevectors (for example, it may include only a part of the known components of the codevector and/or noise). It is multiplied by the matrix of connections with a subsequent binarization by comparison with the threshold. The result is again fed into the input. After several such iterations of evolving the memory, the output codevector is retrieved which is the codevector from the memorized set that is most similar to the input codevector.

We have developed a method for analyzing the characteristics of such a memory [20, 21, 19]. For a wide range of codevector dimensionality, the degree of sparsity, and the level of distortion of codevectors-queries, it was shown that the accuracy of the obtained estimates of information characteristics exceeds the accuracy of the Gibson-Robinson method, and the maximum information efficiency of this memory (per bit of connection implemented in computer memory) is higher than that of the Hopfield network.

Note that the studied learning rule does not allow using this auto-associative memory for generalization. According to Hebb's ideas about cell assemblies, active neurons often found together (corresponding to 1-components of codevectors), when learning by increasing the weight of connections between them, form the "cores" of assemblies, i.e., strongly connected subsets of neurons that fire together more easily. This may correspond, for example, to typical category features and prototype objects. And vice versa, rare combinations of active neurons form a "fringe" corresponding, for example, to features of specific objects, see [19]. To implement this, connections between neurons should be not binary, but gradual, as in the Hopfield network. Another option is to use not a deterministic, but a stochastic learning rule, in which the weight of the connection between 1-components changes with a certain probability, which is set by the "learning rate" parameter [14].

Modeling the formation of cores and fringes in distributed auto-associative memory has so far been investigated only fragmentarily. The same applies to the use of such memory for codevector representations of sequences and structures (but see [16]). These topics are a promising direction for further research [14, 16, 23].

The part-whole hierarchy. In APNNs, the formation of the codevector of an object-whole from the codevectors of its objects-parts or objects-features is performed by superposition or Context-Dependent Thinning operations. A number of questions remains open:

- how to extract objects and their parts of different hierarchical levels?
- how to determine which hierarchical level an object belongs to?
- how to work with an object that can belong to different levels of the hierarchy and modules?

In this regard, it is of great interest to explore possible parallels with how this is done in the brain.

It was shown in [24], that the representation of composite objects is different from the representation of their parts, and there exists a representation in the brain that corresponds to a combination of parts. Moreover, in the process of learning new objects a representation of the object appears which is processed as easily as the representation of a separate feature (so-called unitization).

In [25], a memory model is confirmed in which both the features of the object and the object in the form of bound features are presented. Moreover, the features of an object can be bound not only through their common position, but directly with each other. In addition, it has been found that unbound features can be represented with greater resolution than when they are bound.

In [26], a hierarchy of episode representations is considered, in which the levels of objects, events, and narratives are distinguished. Moreover, the representations of both objects and events also have a hierarchical structure: there exists a representation of both the object-whole and its features [25], and events are represented both by their details as well as by coarser global information. Perhaps there are different mechanisms at work to memorize these different levels of hierarchy. This is manifested in the different nature of forgetting: events are forgotten as a whole, and objects can be forgotten partially, by separate features. In addition, more generalized information is memorized better than details.

Experiments in [27] made it possible to propose an episode recall model in which events can "scroll" forward until the beginning of the next event. In [28], it was experimentally shown that the representation of an object can be bound with both time and position.

The key problem for APNNs is the segmentation of objects (events) into parts and wholes. Progress in its solution can be facilitated by data on the solution of such tasks by the brain. It was shown in [29] that neural states at the upper levels of the hierarchy are activated longer than their parts from the lower levels. This is consistent with the APNN architecture.

When modeling the brain, the representation of spatial structures is usually considered in terms of either cognitive maps or cognitive graphs (see review [30]). In cognitive maps, locations are given by coordinates, and the relationships between them can be viewed in terms of angles and distances, as on a map. In cognitive graphs, only some positions are given by nodes, and the edges between them are path segments, with no information about position or orientation relative to the global coordinate system. Only the topology can be specified (nodes are connected, but the path can be winding) or also labels (distances, directions, angles for existing edges). In [30], it is proposed that representations in the form of cognitive maps and graphs can exist simultaneously, complementing each other, and can also be used to represent not only spatial knowledge. In APNNs, representations of both these data types has been developed [31–33], that can be used both for brain modeling and in practical problems.

Some connected research directions. The representations and basic operations in HDC/VSA provide Turing-completeness. In [34–37], a paradigm different from HDC/VSA that operates with distributed representations is proposed, which is also Turing complete. It is formulated in terms of Hebb assemblies and focuses on the direct handling of assemblies in memory.

To create a "copy" of assembly in a target area (projection operation), a pattern of active neurons is formed in it by randomly projecting the activity pattern of the assembly from the original area and selecting the most active neurons. Note that we considered a similar transformation in [38–41], see also [42]. Then the resulting pattern of active neurons forms a new assembly using variants of the Hebbian rule. This process is complicated by the possibility of modifying the connection weights of a random projection and spreading activity along connections in the target area.

The possibility of back projection is also considered with modification of projection connections from the second area to the first one (bind operation), and merging of assemblies from the two areas to the third one. The last two operations can be used to form assemblies of structures, being analogous to binding and superposition in HDC/VSA. It is interesting to explore how the capabilities of this paradigm relate to the HDC/VSA models, both in terms of applications and biological relevance.

Sketches (see [43–46] and references therein) are compressed representations of data. In [47–48], it was proposed to combine sketches with Deep Neural Networks. Both the formation of sketches of hierarchical structures and the use of memory based on locality-sensitive hashing ([43–46] and references therein) for fast similarity search at each layer of Deep Neural Networks are considered. Random projection is used to form the sketch. The authors of [47–48] consider the problems of recovering input sparse vectors and random matrices themselves, using the projection results. The sparse recovery methods and dictionary learning are used. On the one hand, this aims to overcome the limitations of HDC/VSA related to the lack of learning. On the other hand, this complicates the scheme and introduces additional restrictions related to the formation and handling of sketches.

A theoretical analysis of binding operations other than CDT is considered in [49].

INPUT DATA TRANSFORMATIONS

The key problem in using codevectors to solve practical problems is to obtain them from the input data in such a way that similar codevectors correspond to similar input data (objects). We have developed such transformations for various types of data.

Sets (i.e., collections of elements without specifying an order or other structure) have the simplest codevector representation. Each element of the set is assigned a randomly generated codevector. The codevector of a specific set consisting of specific elements is obtained by superposition (component-wise disjunction). Sets containing the same elements will have similar codevectors. The greater the proportion of identical elements, the greater the similarity.

Numeric data. Numeric data — scalars and vectors — are perhaps the most common data type. We have developed and investigated three types of methods for transforming real vectors a $\in \Re^D$ into codevectors A $\in \{0,1\}^d$.

1. Receptive fields-based methods [45, 46].

The components of a codevector are formed by determining whether the input vector belongs to the receptive fields corresponding to the components of the codevector: $A_i = \psi_i(\mathbf{a})$, $i = 1,...,d$, where A_j are the binary components of the codevector $\mathbf{A}$, $\psi_i(\mathbf{a})$ is the indicator of the vector $\mathbf{a}$ location in some region of the input space, i.e., in the i-th receptive field.

The developed and investigated methods use hyperrectangular receptive fields with random boundaries in random subsets of the dimensions of the input space. This allows a computationally efficient determination of whether a vector belongs to a receptive field by comparison with the boundary values of the field in each of its dimensions. Significant overlap of codevectors of close input vectors is provided due to the presence of a large number of common receptive fields, i.e., due to the vectors falling into a large number of the same receptive fields.

2. Random projection-based methods [43, 44, 46].

The codevector $\mathbf{A}$ is formed by a random linear projection of the input vector $\mathbf{a}$ using a random matrix $\mathbf{R}(d \times D)$ and binarization of the resulting vectors by a component-wise non-linear threshold function T: $\mathbf{A} = T(\mathbf{Ra})$, see Fig. 4. Note that both $d >> D$ and $d << D$ could be the case. Here the matrix $\mathbf{R}$ is a random matrix with elements from a sub-Gaussian distribution. In particular, the Gaussian distribution, ternary one with elements from the set $\{-1, 0, 1\}$, and binary one with elements from $\{0, 1\}$ were studied. Such a transformation can be performed using a perceptron-like neural network with randomly selected connections. Threshold for non-linearity does not have to be zero, and it allows controlling the sparseness of the generated codevectors. Significant overlap of codevectors of close input vectors is provided by a large value of their dot product with the same random vectors — rows of a random matrix. Random projection properties allow estimating the cosine of the angle and the angle of the input vectors from such codevectors. Estimating the Euclidean distance and dot product requires knowledge of the Euclidean norms of the input vectors. A similar neural network architecture was found in the olfactory system of the Drosophila fly [42, 50, 51]. Related problems were also studied in [52, 53].

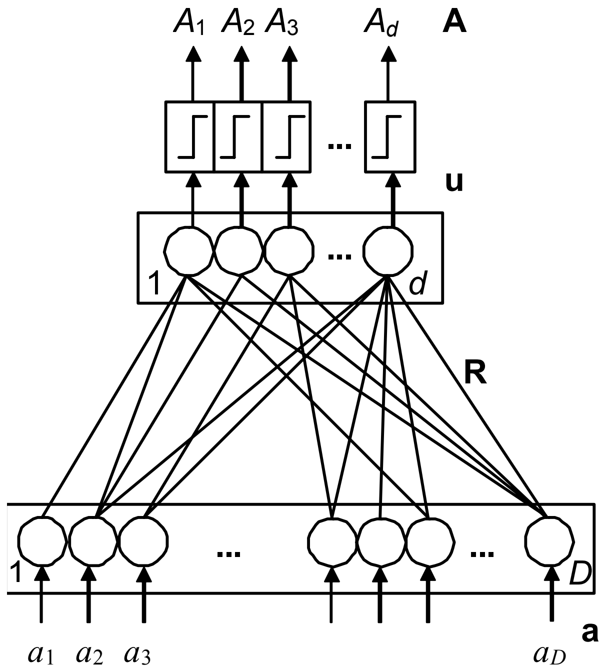


Fig. 4. A single layer perceptron for transformation of vector data by a random projection with a threshold.

3. Compositional methods [44, 45, 54].

For each value of each component-scalar $a(j)$, $j = 1,...,D$, of the input vector **a** with integer components, a codevector is formed, with similar codevectors corresponding to close values of the scalars, and dissimilar ones corresponding to distant values. The codevector **A** of the input **a** is formed from the codevectors $\mathbf{A}_{ja(j)}$ of the values of its components-scalars $a(j)$ как $\mathbf{A} = \langle \mathbf{A}_{1a(1)}, \mathbf{A}_{2a(2)},..., \mathbf{A}_{Da(D)} \rangle$, where $\langle \cdot \rangle$ is the CDT operation. A degenerate case of binding is the component-wise disjunction, in which case, in fact, binding does not occur. The dimension of codevectors for scalars and vectors is the same. Significant overlap of codevectors of close input vectors is ensured by constructing them from similar codevectors of the values of their components.

For some of these methods, the dependence of the probability of coincidence of the components of the codevectors on the value of the components of the input vectors was obtained, providing the similarity function that can be estimated from the dot product of the codevectors.

For all three types of methods, an analysis of the characteristics of codevectors was carried out, which makes it possible to choose their parameters in applications. Note that the non-linearity of the transformation of the input space into codevectors makes it possible to use linear models to solve nonlinear problems using codevectors.

The codevector representation of vector data is used not only in similarity search and classification problems (Sec. 5). We note the use of the method of receptive fields for continuous optimization [55, 56] and the method based on random projection in problems of decentralized flows [57].

Neural regularization approach. Numerical methods used to process vectors and matrices in statistics, machine learning, and inverse problem theory often turn out to be inefficient for large dimensions. This manifests itself both in an increase in computational costs and in the loss of stability of solutions. A productive approach to overcome these problems is the randomization approach. It allows not only to reduce computational costs when searching for solutions, but also, as it turned out, to give stability to numerical methods, see also [58–61].

To improve the stability and accuracy of solving discrete ill-posed inverse problems (DIP), we have developed new methods of neural regularization based on random projections, as well as on the basis of matrix expansions. The methods use an integer regularization parameter that determines the complexity of the linear model. Computable criteria have been developed for choosing the value of the regularization parameter that is optimal from the point of view of the accuracy of solving a discrete ill-posed problem, i.e., of the recovery of an unknown input signal. The application of the developed methods provides not only the stability and accuracy of the solution, but also reduces the computational complexity of the regularization. Our studies of the regularizing properties of random projections started in 2009 [62] and were further developed in [63–67].

An approach and methods for solving DIP based on random projection have been developed. To do this, it was proposed to left-multiply both parts of the approximate equation $\mathbf{Ax} = \mathbf{y}$ by a random matrix $\mathbf{R}_k \in \Re^{k \times N}$ with the number of rows k less than N. The vector of the recovered signal is obtained by multiplying the pseudo-inverse matrix $(\mathbf{R}_k\mathbf{A})^+$ by the right part $\mathbf{R}_k\mathbf{y}$ of the new equation. An experimental study of this basic method showed that averaging over random matrices leads to smoothing of the dependence of input and output recovery errors on k, as well as to a decrease in the number of local minima. As a result, it is easier to find the optimal value of k. This leads to a reduced computational complexity by restricting the search to $k_{\text{opt}+1}$ ($k_{\text{opt}} < N$) values of the criterion that allows getting optimal k (instead of calculating all N values of the criterion). In addition, the error of the recovery of the input vector is reduced relative to the case without averaging.

This gave us reason to believe that analytical averaging over random matrices can give the same useful result. The analytical averaging allowed us to propose a method of the "Deterministic Random Projection" for solving DIP, the error of which is always less than the error of the basic random projection method. Namely, the recovery of the input vector was proposed to be carried out as $\widetilde{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{U}\mathbf{D}_k\mathbf{U}^{\mathrm{T}}\mathbf{y}$ [66], leading to error reduction by the value of variance that appears due to multiplication by a random matrix. Here $\mathbf{D}_k$ is a diagonal matrix of special kind corresponding to $\mathbf{R}_k$, see [66], $\mathbf{U}$ (and $\mathbf{V}$ below) is the matrix of the left (and right) singular vectors.

It was shown that the considered methods of solving DIP (the Tikhonov regularization, the Truncated Singular Value Decomposition, and the Deterministic Random Projection) weigh reciprocal singular values differently when obtaining the solution vector. The expression for estimating the input vector in the general case has the form $\mathbf{x}^* = \mathbf{V}\,diag\left(\frac{1}{s_i} w_i\right)\mathbf{U}^{\mathrm{T}}\mathbf{y}$. For the Tikhonov regularization, it is known that the weights decrease gradually:

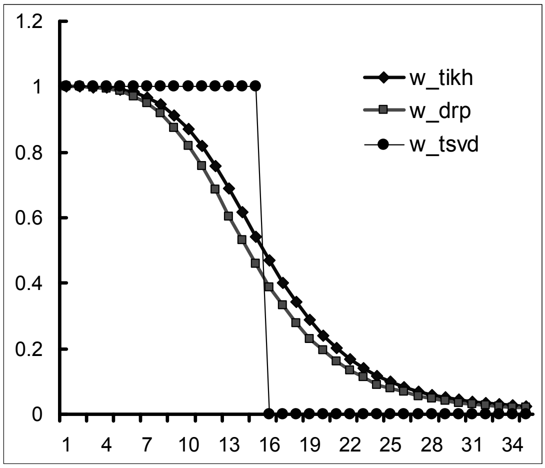


Fig. 5. Examples of the weights of the reciprocal singular values when estimating the input vector by the Tikhonov (w_tikh), the Deterministic Random Projection (w_drp), and the Truncated Singular Value Decomposition (w_tsvd) types of regularizations.

$$\mathbf{x}_{\text{Tikh}} = (\mathbf{A}^{\text{T}}\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}^{\text{T}}\mathbf{y} = \mathbf{V}diag\left(\frac{s_i^2}{(s_i^2+\lambda)}\frac{1}{s_i}\right)\mathbf{U}^{\text{T}}\mathbf{y},\ w_{\text{Tikh}} = \frac{s_i^2}{(s_i^2+\lambda)}.$$

For the Truncated Singular Value Decomposition, the weights of $1/s_1 \ldots 1/s_k$ are equal to 1, whereas the weights for $1/s_{k+1}\ldots 1/s_N$ are equal to 0. For the Deterministic Random Projection,

$$\mathbf{x}_{\text{DRP}} = \mathbf{A}^{\text{T}}\mathbf{U}\mathbf{D}_k\mathbf{U}^{\text{T}}\mathbf{y} = \mathbf{V}\mathbf{S}\mathbf{D}_k\mathbf{U}^{\text{T}}\mathbf{y} = \mathbf{V}diag\left(s_i^2 d_{ki}\frac{1}{s_i}\right)\mathbf{U}^{\text{T}}\mathbf{y},\ w_{\text{DRP}} = s_i^2 d_{ki}.$$

Here d_{ki} are the diagonal elements of $\mathbf{D}_k$. So, the values of the weights gradually decrease with the increasing index of the singular value, see Fig. 5. This is similar to the behavior of weights for the Tikhonov regularization and provides a potentially higher accuracy compared to the Truncated Singular Value Decomposition, due to a better signal approximation.

Structured data. A number of methods for codevector representation of data with the sequence structure have been developed in [68, 69, 31]. They are based on the formation of codevectors which represent the elements of a sequence according to their positions. These codevectors are combined into a codevector of the whole sequence either by a superposition operation (component-wise disjunction) or by a binding operation. A widespread variety of sequences are strings, where the elements are symbols at sequential positions.

The following approaches have been developed to represent sequence elements taking into account their position:

1) multiplicative binding of codevectors of sequence elements with codevectors of their positions, as well as with codevectors of context elements [70, 16];

2) the use of permutations of codevectors of sequence elements to represent their order [71, 12];

3) representation by means of codevectors of n-grams (i.e., n consecutive elements) of a sequence [69, 72].

In the first approach, to represent identical or similar sequence elements at close positions with similar codevectors, correlated codevectors have been proposed as codevectors for close positions. In the second approach, to preserve the similarity of element codevectors at close positions, it was proposed in [73] to use partial (correlated) permutations. In the third approach, the representation of n-grams and whole strings by multiplicative binding of symbol codevectors permuted according to their position was proposed in [74]. In [69], it was considered how to approximate the edit distance of strings by vectors of frequencies of different n-grams, allowing a codevector implementation.

The recently proposed methods [31, 75] made it possible to form similar codevectors for sequences with similar elements at close positions. These methods, in contrast to the previous ones, allow obtaining codevectors of transformed (shifted) sequences not only by their formation from the transformed sequences, but also by transforming the codevector of the original sequence (the equivariance property).

In addition to the 1D case of sequences, several approaches have been developed for the codevector representation of objects with a two-dimensional structure, such as 2D images.

In the role-filler approach, codevectors of positions are formed in such a way that similar codevectors correspond to close positions. The features extracted from the images are also associated with codevectors. A feature at its position is represented as a codevector obtained by binding the codevectors of the feature and its position. The codevector of the entire image is obtained by superposition of the codevectors of all extracted features at their positions. Such codevector representations are similar for images that contain similar features at similar positions.

In the permutation-based approach, the representation of the feature codevector at its position is performed by permutations, and the resulting codevectors are superimposed to represent the entire image. The use of partial permutations makes it possible to ensure the similarity of codevectors of a feature at close positions [73].

Arbitrary binary features can be used as features. So, for the direct formation of codevectors from images, special binary LIRA features were proposed [76, 77]. For binary images, each LIRA feature is an indicator of the presence of 1/0 pixels at randomly selected positions of a randomly located local window. For gradual images, the brightness of pixels is compared against randomly selected thresholds. Each LIRA feature corresponds to a codevector component. The parameters are chosen so that the image usually contains a small fraction of the entire set of features, i.e., codevector is sparse.

The RLD features (Random Local Descriptors) can be considered as the development of LIRA features. In the RLD approach [73], the same set of features is extracted at each “interesting” point in the image. Each feature is assigned a random codevector. Features at their positions are represented by partial permutations of the corresponding codevectors. As a result, features at close positions produce similar codevectors. Due to this, RLD improves the results of LIRA in classification problems. Currently, methods for generating image codevectors are being developed that provide equivariance to some image transformations.

Complex structured data, such as hierarchically organized graphs of knowledge base episodes, are initially described by systems of hierarchically organized objects and relations. Based on the approaches from the previous sections, a number of methods for codevector representation of such data have been developed, their basic blocks being codevector representations of relations. The codevectors of relations are recursively transformed into codevectors of complex hierarchical structures containing higher-order relations [10, 11, 18, 32, 33]. The codevectors of structures have the same dimension as the codevectors of their elements.

Consider an example of the Solar System episode representation (please see details in [33]) shown in a bracketed notation in Fig. 6 and as a graph sketch in Fig. 7. It includes objects Sun, Planet; attributes Mass, Temperature; relations Gravity, Attracts, Greater, Revolve-Around, And; higher-order relations Cause. Relations have arguments; the relation's name together with the particular arguments form the instance of a relation. For many types of relations, the order of arguments is essential. So, a relational instance can be described by roles specific to the relation and instances of arguments that fill them. To form the codevector of a relational instance, the role-filler approach uses a multiplicative binding of the role and filler codevectors to indicate which role the filler is assigned to in a relation (e.g., the second role in Greater is filled with a smaller object instance). The codevector of the whole episode is formed as shown in Fig. 8. In the predicate-arguments approach, random permutations of the codevectors of the arguments of a relation are used to represent their order, as in the representation of sequences.

```
(CAUSE
        (GRAVITY (MASS SUN) (MASS PLANET))
        (ATTRACTS SUN PLANET) )
(GREATER ( TEMPERATURE SUN)
        ( TEMPERATURE PLANET) )
(CAUSE
        (AND  (GREATER (MASS SUN)
                        (MASS PLANET) )
               (ATTRACTS SUN PLANET) )
        (REVOLVE-AROUND PLANET SUN) )
```

Fig. 6. Bracketed notation description of the Solar System episode

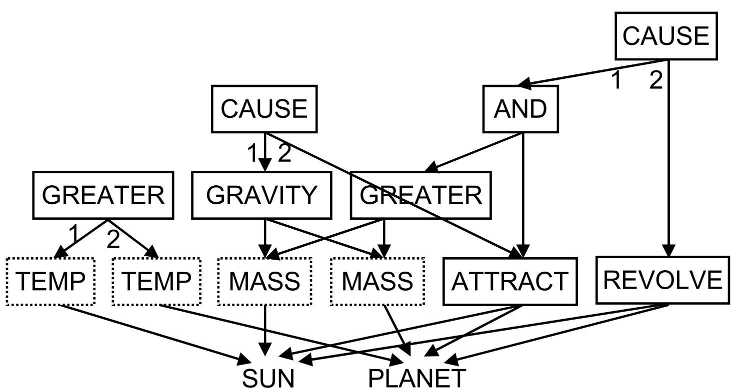


Fig. 7. Graph description of the Solar System episode.

```
SOLAR_SYSTEM =
⟨ CAUSE_1 ∨ ⟨GRAVITY_1 ∨ ⟨MASS ∨ SUN⟩ ⟩ ∨ ⟨GRAVITY_2 ∨ ⟨MASS ∨ PLANET⟩⟩ ⟩
∨ ⟨CAUSE_2 ∨ ⟨ATTRACTS_1 ∨ SUN⟩ ∨ ⟨ATTRACTS_2 ∨ PLANET⟩ ⟩
∨
⟨ GREATER_1 ∨ ⟨TEMPERATURE ∨ SUN ⟩⟩
∨ ⟨ GREATER_2 ∨ ⟨TEMPERATURE ∨ PLANET⟩ ⟩
∨
⟨ CAUSE_1 ∨
   ⟨ AND ∨ ⟨GREATER_1∨⟨MASS∨SUN⟩⟩ ∨ ⟨GREATER_2∨⟨MASS∨PLANET⟩⟩ ⟩
  ∨ ⟨ AND ∨ ⟨ATTRACTS_1 ∨ SUN⟩ ∨ ⟨ATTRACTS_2 ∨ PLANET⟩⟩ ⟩
∨ ⟨ CAUSE_2 ∨
   ⟨ REVOLVE-AROUND_1 ∨ PLANET⟩ ∨ ⟨REVOLVE-AROUND_2∨SUN⟩ ⟩
```

Fig. 8. Codevector representation of the Solar System episode.

The similarity of codevectors of similar complex structured data is ensured by the similarity of the codevectors of their elements and the properties of binding operations. As a result, for structures with similar objects and relations, the created methods ensure the production of similar codevectors. Thus, by finding the similarity of the codevectors of relational structures using some vector similarity measure of vectors (for example, the dot product), we simultaneously evaluate the similarity of structures and the similarity of objects in these structures. This provides the basis for the creation of computationally efficient and qualitatively new methods for processing relational structures of data- and knowledge bases, which are based on similarity and simultaneously take into account both the structure and semantics of knowledge.

METHODS

Based on the proposed approaches, methods for solving various types of problems from the field of Machine Learning and Artificial Intelligence have been developed.

The approach of "example-based reasoning" [45, 46] is productively used by humans when solving problems of natural intelligence; it is also used for solving a wide range of problems in Artificial Intelligence systems. For inferences about a query (an input object or a situation), this approach uses a search for similar known objects or situations with which additional information is associated. In computer implementation, a base of objects-example is formed, i.e., a memory containing the past "experience" of the system. Examples found by similarity search in memory can be used directly, or as a source of additional information about the input object-query. One example of using this approach to solve classification problems is the nearest neighbor method, where the class label of the nearest (by some measure of similarity) example from the memory is transferred to the query. Another example is a linear classifier, where for each class, in the process of learning a vector representation of the generalized example is formed, and the classification is performed by choosing the class whose generalized example gives the maximum value of the dot product with the vector representation of the query.

The effectiveness of applying the example-based reasoning approach to solving problems depends on how the similarity between examples is defined. When using codevectors, this, in turn, depends on the methods of their formation and the applied similarity measures of codevectors. We also note the existence of efficient algorithms for fast similarity search [78, 79, 46].

The developed methods of codevector representation of data of various types make it possible to solve classification problems using linear vector classifiers. In particular, for vector data, often the class boundaries are not linearly separable. However, non-linear data transformation to codevectors overcomes this problem. Combining a specific transformation of input data into codevectors with a specific type of linear classifier yields a new version of the classifier [80–84]. The best-known linear classifiers are perceptrons and Support Vector Machines. We have proposed a perceptron with a large margin that approximates the results of Support Vector Machine, but is trained incrementally and fast [84]. As mentioned above, other types of classifiers can be used, such as the nearest neighbor classifier.

The architecture of a modular neural network with an assembly organization has been developed that can be considered as a generalization of the perceptron classifier [83]. Each module corresponds to a class. In the fully connected version, each neuron is connected by trainable connections with all other neurons of the module, and in the non-fully connected version, with a randomly selected subset of neurons. Learning is performed similarly to the perceptron learning rule, i.e., by increasing the weights of connections between active neurons in the module corresponding to the correct class, and decreasing the weights of connections in the module of the wrong class. In the recognition mode, neural activity is propagated along the connections in each module, and the activity of the module is calculated as the total sum activity of the module's neurons. The module with the highest activity determines the winning class.

Also, the architecture of a layered neural network with competitive layers for image processing has been developed. Each layer corresponds to a certain class of images and is a separate neural network. Neurons from different layers have a one-to-one correspondence with each other and with the 2D input retina. There is a competition between the corresponding neurons of all layers resulting in the activation of single most active neuron among all layers in each retina's position. Such networks have been applied to the problems of classifying handwritten digits, texture segmentation, and extracting image segments of different orientations [85, 86].

Concerning texture segmentation, a method has been developed for segmenting visual images into homogeneous regions of fine-grained texture [87, 88]. The peculiarities of the method are that it works without training, and no preliminary information about the analyzed image is required.

Work is underway to use the developed methods in tasks related to Unmanned Aerial Vehicles [89, 90].

APPLICATIONS

The effectiveness of the developed methods has been demonstrated by solving a wide range of applied problems. Upon the time of the original publications, a number of the results obtained were comparable to the state-of-the-art as indicated below.

To take into account the semantics of textual information (words, texts and their fragments) presented in the form of frequency vectors of the joint occurrence of words and their contexts, methods for the formation of "context codevectors" have been developed. In the task of searching for textual information, due to taking into account semantics, the accuracy of the search was increased up to 20% on the text datasets Time, Cranfield, Medlars [91]. Context codevectors have also been applied in tasks that require taking into account the semantic similarity of words, including the search for synonyms and the choice of the proper target word in automatic translation.

Strong performance has been obtained in solving problems of classification and segmentation of textures [15, 86], recognition of handwritten characters and words [84, 92], acoustic signals [81, 93], etc. In text classification, on the Reuters-21578 dataset the accuracy was increased to 0.937 corresponding to the best contemporary results. In predicting the sensitivity of a cancerous tumor of glioma to chemotherapy, the prediction accuracy was improved to 88.5% compared to less than 81% for other methods. A classification accuracy of up to 99.5% has been achieved on the MNIST handwritten digits dataset [73, 76, 77, 84, 92]. It was shown that for a small feature set, a non-fully connected modular network improved the results of the perceptron classifier.

A high-precision system for recognizing a speaker by voice in noisy conditions has been created [93]. The individual features of the voices were fixed in the structure of the neural network in the process of training neural network classifiers. The network automatically generated individual portraits of voices as a collection of speech features. The system worked both in search mode in voice databases and in real time. The reliability of identification of microphone signals was within 94%–98% and reached 85%–94% for phone signals. The technology worked with the datasets that included an arbitrary number of voice samples from various people recorded via microphone and phone channels. The ability to search for defined voices in many hours of audio recordings was also implemented (audio data indexing).

Codevector representations of data with the structure of sequences have been tested in the tasks of detecting spam and intrusions in computer systems, classifying coding regions of genes, predicting the structure of proteins, searching for text duplicates, spellchecking, and modeling the visual similarity of words in humans [69, 31].

The developed methods for generating codevectors of complex structured data that include relations (Sec. 4.3) were used in reasoning by analogy to effectively search for analogs by similarity while simultaneously taking into account structure and semantics. The application of the proposed approach made it possible, on the ThinkNet knowledge base, to increase the search precision and recall up to 20% and 4 times, correspondingly [33]. In addition, methods for analogical mapping and inference were developed and tested.

In the problem of predicting the existence of chemical compounds on the INTAS00-397 dataset, the obtained results [94] exceeded the best known ones.

Due to the new methods of neural network regularization, systems with increased accuracy have been developed for gamma spectrometry at fixed and non-fixed measurement geometries [95], suppression of active interference [96], estimating the direction of signal arrival in antenna systems [97].

CONCLUSIONS AND PERSPECTIVES

The key issue in the problems of Artificial Intelligence is the adequate representation of data. The approach under development is based on the idea of distributed representation of information in the brain and allows representing various types of data, from numeric values to graphs, as vectors of large but fixed dimensionality. The similarity of the initial data is manifested in the similarity of the resulting vectors. This makes it possible to apply similarity search in solving a number of problems based on case-based reasoning, presumably similar to how the human brain does, and also allows using an extensive arsenal of existing vector machine learning methods for processing and analysis.

We also develop theoretically substantiated feature extraction methods based on sparse multilayer neural network models and new approaches to regularization. The methods are applicable to data that allow both 1D and 2D representation, such as sequences (audio signals), 2D images, video sequence frames, and so on.

Recently, well-known figures in the field of Deep Neural Networks, such as G. Hinton [98], J. Bengio [99], Y. Schmidhuber [100] proposed that new ideas are required to overcome the shortcomings of Deep Neural Networks. The sense of the new direction of research is to provide the ability to form and operate structures consisting of internal representations of objects, without learning such representations from scratch. The methods being developed by us are also aimed to achieve such properties of "compositionality".

The main area of research reviewed in this paper is the problem of representing heterogeneous data in a unified format for the Artificial Intelligence systems based on modeling the neural network organization of the brain and the mechanisms of thinking hypothesized by Amosov. The most important advantages of the developed approach are the possibility of natural integration and efficient processing of various types of data and knowledge, a high degree of parallel computing, reliability and resistance to noise, the possibility of hardware implementation with high performance and energy efficiency, data processing based on associative search by similarity, similar to how human memory works. This allows one to unify methods, algorithms, software, and hardware for Artificial Intelligence systems, to increase their scalability in terms of speed and memory with an increase in data volume and complexity.

Currently, the topical direction of our developments is the creation of vector distributed representations of objects that would allow their modification inside the Artificial Intelligence system so that the result coincides with the representation of the external objects after some (corresponding) transformations. This applies to techniques for equivariantly representing sequences, as well as spatial objects such as 2D images and higher dimensional representations, including making such representations equivariant to translation, rotation, and scale. Such representations may be considered as analogous to imagery in human thinking, and operating with them may be seen as a form of creative thinking.

Another promising direction is the consideration of context. Many brain experiments show that the context has a great influence on the memorization of objects, events, scenes, etc. In APNNs, context is taken into account by the binding of codevectors that is performed by Context-Dependent Thinning, as well as by permutation. It is interesting to test the hypothesis that binding an object's codevector and a particular context's codevector yields strongly context-dependent representations of objects that do not allow the object to be recognized in another context. However, when an object is stored in many different contexts, a context-independent representation of the object is formed.

We believe that the reviewed studies create the basis for overcoming some of the shortcomings of modern approaches to specialized Artificial Intelligence based on Deep Neural Networks and will contribute to the development of Artificial General Intelligence.

REFERENCES

1. Amosov N.M. Modelling of thinking and the mind. New York: Spartan Books, 1967, 192 p.
2. Amosov N.M., Kasatkin A.M., Kasatkina L.M., Kussul E.M., Talaev S.A. Intelligent behaviour systems based on semantic networks. *Kybernetes*. 1973, Vol. 2, N 4, pp. 211–216.
3. Amosov N.M., Kussul E.M., Fomenko V.D. Transport robot with a neural network control system. *Advance papers of 4th Intern. Joint Conf. on Artif. Intelligence*. 1975, Vol 9, pp. 1–10.
4. Amosov N.M. Algorithms of the Mind. Kiev: Naukova Dumka, 1979, 223 p. (in Russian)
5. Amosov N.M., Baidyk T.N., Goltsev A.D., Kasatkin A.M., Kasatkina L.M., Rachkovskij D.A. Neurocomputers and intelligent robots. Kiev: Naukova Dumka. 1991, 269 p. (in Russian)
6. Kussul E.M. Associative neuron-like structures. Kiev: Naukova Dumka. 1992, 144 p. (in Russian)
7. Kleyko D., Rachkovskij D.A., Osipov E., Rahimi A. A survey on Hyperdimensional Computing aka Vector Symbolic Architectures, Part I: Models and data transformations. *ACM Computing Surveys*. 2022. https://doi.org/10.1145/3538531
8. Kleyko D., Rachkovskij D.A., Osipov E., Rahimi A. A survey on Hyperdimensional Computing aka Vector Symbolic Architectures, Part II: Applications, Cognitive Models, and Challenges. Accepted, *ACM Computing Surveys*. 2022. Available online: *arXiv:2112.15424.*
9. Thorpe S. Localized versus distributed representations The Handbook of Brain Theory and Neural Networks. Edited by M.A. Arbib.Cambridge, MA: The MIT Press. 2003, pp. 643–646.
10. Rachkovskij D.A., Kussul E.M. Binding and normalization of binary sparse distributed representations by context-dependent thinning. *Neural Computation*. 2001, Vol. 13, N 2, pp. 411–452.
11. Plate T. Holographic Reduced Representation: Distributed Representation for Cognitive Structures. Stanford: CSLI Publications, 2003, 300 p.
12. Kanerva P. Hyperdimensional computing: an introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*. 2009, Vol. 1, N2, pp. 139–159.
13. Gayler R.W. Multiplicative binding, representation operators, and analogy. Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences. Sofia, Bulgaria: New Bulgarian University, 1998, p. 405.
14. Kussul E.M., Rachkovskij D.A., Baidyk T.N. Associative-Projective Neural Networks: Architecture, Implementation, Applications. *Proc. Neuro-Nimes'91*. 1991, pp. 463–476.
15. Kussul E.M., Rachkovskij D.A., Baidyk T.N. On image texture recognition by associative-projective neurocomputer. *Proc. ANNIE'91 Conference "Intelligent engineering systems through artificial neural networks"*. St. Louis, MO: ASME Press, 1991, pp. 453–458.
16. Kussul E.M., Rachkovskij D.A. Multilevel assembly neural architecture and processing of sequences. In Neurocomputers and Attention: Connectionism and Neurocomputers, vol. 2. Manchester and New York: Manchester University Press, 1991, pp. 577–590.
17. Gritsenko V.I., Rachkovskij D.A., Goltsev A.D., Lukovych V.V., Misuno I.S., Revunova E.G., Slipchenko S.V., Sokolov A.M., Talayev S.A. Neural distributed representation for intelligent information technologies and modeling of thinking. *Cybernetics and Computer Engineering*. 2013, Vol. 173, pp. 7–24. (in Russian).
18. Rachkovskij D.A., Kussul E.M., Baidyk T.N. Building a world model with structure-sensitive sparse binary distributed representations. *Biologically Inspired Cognitive Architectures*. 2013, Vol. 3, pp. 64–86.
19. Gritsenko V.I., Rachkovskij D.A., Frolov A.A., Gayler R., Kleyko D., Osipov E. Neural distributed autoassociative memories: A survey. *Cybernetics and Computer Engineering*. 2017, N 2 (188), pp. 5–35
20. Frolov A.A., Rachkovskij D.A., Husek D. On information characteristics of Willshaw-like auto-associative memory. *Neural Network World*. 2002. Vol. 12, No 2, pp. 141–158.
21. Frolov A.A., Husek D., Rachkovskij D.A. Time of searching for similar binary vectors in associative memory. *Cybernetics and Systems Analysis*. 2006, Vol. 42, N 5, pp. 615–623.
22. Fusi S. Memory capacity of neural network models. *arXiv:2108.07839*. 21 Dec 2021.
23. Steinberg J., Sompolinsky H. Associative memory of structured knowledge. 2022. https://doi.org/10.1101/2022.02.22.481380

24. Liang J.C., Erez J., Zhang F., Cusack R., Barense M.D. Experience transforms conjunctive object representations: Neural evidence for unitization after visual expertise. *Cerebral Cortex*. 2020. Vol. 30, N 5, pp. 2721–2739.
25. Li A.Y., Fukuda K., Barense M.D. Independent features form integrated objects: Using a novel shape-color “conjunction task” to reconstruct memory resolution for multiple object features simultaneously. *Cognition*. 2022, Vol. 223, Article 105024.
26. Andermane N., Joensen B.H., Horner A.J. Forgetting across a hierarchy of episodic representations. *Current Opinion in Neurobiology*. 2021, Vol. 67, pp. 50–57.
27. Michelmann S., Hasson U., Norman K.A. Event boundaries are steppingstones for memory retrieval. 2021. Preprint DOI 10.31234/osf.io/k8j94.
28. Schneegans S., McMaster J.M.V., Bays P.M. Role of time in binding features in visual working memory. *Psychological Review.* 2022. https://doi.org/10.1037/rev0000331
29. Geerligs L., van Gerven M., Campbell K.L., Güçlü U. A nested cortical hierarchy of neural states underlies event segmentation in the human brain. *Neuroscience.* 2021. https://doi.org/10.1101/2021.02.05.429165
30. Peer M., Brunec I.K., Newcombe N.S., Epstein R.A. Structuring knowledge with cognitive maps and cognitive graphs. *Trends in Cognitive Sciences*. 2021, Vol. 25, N 1, pp. 37–54.
31. Rachkovskij D.A. Shift-equivariant similarity-preserving hypervector representations of sequences. *arXiv:2112.15475*. 2021.
32. Rachkovskij D.A. Representation and processing of structures with binary sparse distributed codes. *IEEE Trans. Knowledge Data Engineering*. 2001, Vol. 13, N 2, pp. 261–276.
33. Rachkovskij D.A., Slipchenko S.V. Similarity-based retrieval with structure-sensitive sparse binary distributed representations. *Comput. Intelligence*. 2012,Vol. 28, N. 1, pp. 106–129.
34. Papadimitriou C.H., Vempala S.S., Mitropolsky D., Collins M.J., Maass W. Brain computation by assemblies of neurons. *Proceedings of the National Academy of Sciences*. 2020, Vol. 117, N 25, pp. 14464–14472.
35. Müller M.G., Papadimitriou C.H., Maass W., Legenstein R. A model for structured information representation in neural networks of the brain. *eNeuro*. 2020, Vol. 7, N 3, pp. 1–17.
36. Mitropolsky D., Collins M.J., Papadimitriou C.H. A biologically plausible parser. *Transactions of the Association for Computational Linguistics*. 2021, Vol. 9, pp. 1374–1388.
37. Papadimitriou C.H., Friederici A.D. Bridging the gap between neurons and cognition through assemblies of neurons. *Neural Computation*. 2022, Vol. 34, N 2, pp. 291–306.
38. Rachkovskij D.A., Misuno I.S., Slipchenko S.V. Randomized projective methods for the construction of binary sparse vector representations. *Cybernetics and Systems Analysis*. 2012, Vol 48, N 1, pp. 146–156.
39. Rachkovskij D.A. Vector data transformation using random binary matrices. *Cybernetics and Systems Analysis*. 2014, Vol. 50, N 6, pp. 960–968.
40. Rachkovskij D.A. Formation of similarity-reflecting binary vectors with random binary projections. *Cybernetics and Systems Analysis*. 2015, Vol. 51, N 2, pp. 313–323.
41. Rachkovskij D.A. Estimation of vectors similarity by their randomized binary projections. *Cybernetics and Systems Analysis*. 2015, Vol. 51, N 5, pp. 808–818.
42. Dasgupta S., Stevens C.F., Navlakha S. A neural algorithm for a fundamental computing problem. *Science*. 2017, Vol. 358, pp. 793–796.
43. Rachkovskij D.A., Gritsenko V.I. Distributed Representation of Vector Data Based on Random Projections. Kyiv: Interservice, 2018. ISBN: 978-617-696-837-5. (in Ukrainian)
44. Rachkovskij D.A. Codevectors: Sparse Binary Distributed Representations of Numerical Data. Kiev: Interservice, 2019. ISBN: 978-617-696-987-7. (in Russian)
45. Gritsenko V.I., Rachkovskij D.A. Methods for Vector Representation of Objects for Fast Similarity Estimation. Kyiv: Naukova Dumka, 2019. ISBN: 978-966-00-1741-2. (In Russian)
46. Rachkovskij D. A. Introduction to fast similarity search. Kyiv: Interservice, 2019, 294 p. ISBN: 978-617-696-904-4. (in Russian)
47. Ghazi B., Panigrahy R., Wang J. Recursive sketches for modular deep learning. *Proceedings of the 36th International Conference on Machine Learning*. *PMLR*. 2019, Vol. 97, pp. 2211–2220.
48. Panigrahy R., Wang X., Zaheer M. Sketch based memory for neural networks. *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics* (*AISTATS'21*). 2021, San Diego, California, USA. *PMLR*. 2021, Vol. 130, pp. 3169–3177.

49. Hiratani N., Sompolinsky H. Optimal quadratic binding for relational reasoning in vector symbolic neural architectures. *arXiv:2204.07186*. 2022, pp. 1–32.
50. Chaitanya R., Hopfield J., Grinberg L., Krotov D. Bio-inspired hashing for unsupervised similarity search. *Proceedings of the 37th International Conference on Machine Learning. PMLR*. 2020, Vol. 119, pp. 8295–8306.
51. Liang Y., Ryali C., Hoover B., Grinberg L., Navlakha S., Zaki M., Krotov D. Can a fruit fly learn word embeddings? *Proc. ICLR'21*. 2021.
52. Li W. Modeling winner-take-all competition in sparse binary projections. In: Machine Learning and Knowledge Discovery in Databases. Edited by: F. Hutter, K. Kersting, J. Lijffijt, I. Valera. *Lecture Notes in Computer Science*. Vol. 12457. Cham: Springer, 2021, pp. 456–472. https://doi.org/10.1007/978-3-030-67658-2_26
53. Li W.Y., Zhang S.Z. Binary random projections with controllable sparsity patterns. *Journal of the Operations Research Society of China*. 2022. https://doi.org/10.1007/s40305-021-00387-0
54. Rachkovskij D.A., Slipchenko S.V., Misuno I.S., Kussul E.M., Baidyk T.N. Sparse binary distributed encoding of numeric vectors. *Journal of Automation and Information Sciences*. 2005, Vol. 37, N 11, pp. 47–61.
55. Izawa S., Kitai K., Tanaka S., Tamura R., Tsuda K. Continuous black-box optimization with quantum annealing and random subspace coding. *Proc. Adiabatic Quantum Computing* (AQC'21), June 22-25, 2021.
56. Izawa S., Kitai K., Tanaka S., Tamura R., Tsuda K. Continuous black-box optimization with an Ising machine and random subspace coding. *Phys. Rev. Research*. 2022, Vol. 4, N 2. Article 023062.
57. Barclay I. et al. Trustable service discovery for highly dynamic decentralized workflows. *Future Generation Computer Systems*. 2022. DOI: 10.1016/j.future.2022.03.035
58. Wei Y., Xie P., Zhang L. Tikhonov regularization and randomized GSVD. *SIAM J. Matrix Analysis Appl.* 2016, Vol. 37, pp. 649–675.
59. Zhang L., Wei Y. Randomized core reduction for discrete ill-posed problem. *J. Comput. Appl. Math*. 2020, Vol. 375, Article 112797.
60. Wei W., Zhang H., Yang X., Chen X. Randomized generalized singular value decomposition. *Commun. Appl. Math. Comput.* 2021, Vol 3, pp. 137–156.
61. Zuo Q., Wei Y., Xiang H. Quantum-inspired algorithm for truncated total least squares solution. *arXiv:2205.00455*. 2022
62. Revunova E.G., Rachkovskij D.A. Using randomized algorithms for solving discrete ill-posed problems. *J. Information Theories and Applications*. 2009, Vol. 16, N 2, pp. 176–192.
63. Rachkovskij D.A., Revunova E.G. Randomized method for solving discrete ill-posed problems. *Cybernetics and Systems Analysis*. 2012, Vol. 48, N 4, pp. 621–635.
64. Revunova E.G. Analytical study of the error components for the solution of discrete ill-posed problems using random projections. *Cybernetics and Systems Analysis*. 2015, Vol. 51, N 6, pp. 978–991.
65. Revunova E.G. Model selection criteria for a linear model to solve discrete ill-posed problems on the basis of singular decomposition and random projection. *Cybernetics and Systems Analysis*. 2016, Vol. 52, N 4, pp. 647–664.
66. Revunova E.G. Increasing the accuracy of solving discrete ill-posed problems by the random projection method. *Cybernetics and Systems Analysis*. 2018, Vol. 54, N 5, pp. 842–852.
67. Revunova O.G., Tyshchuk O.V., Desiateryk O.O. On the generalization of the random projection method for problems of the recovery of object signal described by models of convolution type. *Control Systems and Computers*. 2021, N 5–6, pp. 25–34.
68. Sokolov A., Rachkovskij D. Approaches to sequence similarity representation. *Int. J. Inf. Theor. Appl*. 2006, Vol. 13, N 3, pp. 272–278.
69. Sokolov A. Vector representations for efficient comparison and search for similar strings. *Cybernetics and Systems Analysis*. 2007, Vol. 43, N 4, pp. 484–498.
70. Rachkovskij D.A. Development and investigation of multilevel assembly neural networks. PhD thesis. Kiev, Ukrainian SSR: V.M. Glushkov Institute of Cybernetics, 1990. (in Russian)
71. Kussul E.M., Baidyk T.N. On information encoding in Associative-Projective Neural Networks. Technical Report 93-3. V.M. Glushkov Institute of Cybernetics, 1993. (in Russian)

72. Kleyko D., Osipov E., De Silva D., Wiklund U., Vyatkin V., Alahakoon D. Distributed representation of n-gram statistics for boosting self-organizing maps with hyperdimensional computing. *Proc. PSI'19.* 2019, pp. 64–79.
73. Kussul E.M., Baidyk T.N., Wunsch D.C., Makeyev O., Martin A. Permutation coding technique for image recognition system," *IEEE Trans. Neural Networks*. 2006, Vol. 17, N 6, pp. 1566–1579.
74. Rachkovskij D.A. Application of stochastic assembly neural networks in the problem of interesting text selection. *Neural network systems for information processing*. 1996, pp. 52–64. (in Russian)
75. Rachkovskij D.A., Kleyko D. Recursive binding for similarity-preserving hypervector representations of sequences. *Proc. IJCNN'22.* 2022.
76. Kussul E., Baidyk T. Improved method of handwritten digit recognition tested on MNIST database. *Image and Vision Computing*. 2004, Vol. 22, pp. 971–981.
77. Makeyev O., Sazonov E., Baidyk T., Martin A. Limited receptive area neural classifier for texture recognition of mechanically treated metal surfaces. *Neurocomputing*. 2008, Vol. 71, N 7-9, pp. 1413–1421.
78. Rachkovskij D.A. Distance-based index structures for fast similarity search. *Cybernetics and Systems Analysis*. 2017, Vol. 53, N 4, pp. 636–658.
79. Rachkovskij D.A. Index structures for fast similarity search for binary vectors. *Cybernetics and Systems Analysis*. 2017, Vol. 53, N 5, pp. 799–820.
80. Kussul E.M., Baidyk T.N., Lukovich V.V., Rachkovskij D.A. Adaptive neural network classifier with multifloat input coding. *Proc. NeuroNimes'93*, Nimes, France, Oct. 25-29, 1993, pp. 209–216.
81. Lukovich V.V., Goltsev A.D., Rachkovskij D.A. Neural network classifiers for micromechanical equipment diagnostics and micromechanical product quality inspection. *Proc. EUFIT'97*, 1997, pp. 534–536.
82. Rachkovskij D.A., Kussul E.M. DataGen: a generator of datasets for evaluation of classification algorithms. *Pattern Recognition Letters*. 1998, Vol.19, N 7, pp. 537–544.
83. Goltsev A.D. Neural networks with assembly organization. Kyiv: Naukova Dumka. 2005, 200 p. (in Russian)
84. Rachkovskij D. Linear classifiers based on binary distributed representations. *J. Inf. Theories Appl.* 2007, Vol. 14, N 3, pp. 270–274.
85. Goltsev A., Rachkovskij D. A recurrent neural network for partitioning of hand drawn characters into strokes of different orientations. *International Journal of Neural Systems*. 2001, Vol. 11, pp. 463–475.
86. Goltsev A., Gritsenko V., Húsek D. Segmentation of visual images by sequential extracting homogeneous texture areas. *Journal of Signal and Information Processing*. 2020, Vol. 11, N 4, pp. 75–102.
87. Goltsev A. D., Gritsenko V.I. Algorithm of sequential finding the textural features characterizing homogeneous texture segments for the image segmentation task. *Cybernetics and Computer Engineering*. 2013, N 173, pp. 25–34.
88. Goltsev A., Gritsenko V., Kussul E., Baidyk T. Finding the texture features characterizing the most homogeneous texture segment in the image. *Proc. IWANN'15*. 2015, pp. 287–300.
89. Volkov O., Komar M., Volosheniuk D. Devising an image processing method for transport infrastructure monitoring systems. *Eastern-European Journal of Enterprise Technologies*. 2021, Vol. 4, N 2 (112), pp. 18–25. https://doi.org/10.15587/1729-4061.2021.239084
90. Gritsenko V., Volkov O., Komar M., Volosheniuk D. Integral adaptive autopilot for an unmanned aerial vehicle. *Aviation*. 2018, Vol. 22, N 4, pp. 129–135. http://dx.doi.org/10.3846/aviation.2018.6413
91. Misuno I. S., Rachkovskij D. A., Slipchenko S. V., Sokolov A. M. Searching for text information with the help of vector representations. *Probl. Programming*. 2005, N 4, pp. 50–59.
92. Goltsev A., Rachkovskij D. Combination of the assembly neural network with a perceptron for recognition of handwritten digits arranged in numeral strings. *Pattern Recognition*. 2005, Vol. 38, N 3, pp. 315–322.
93. Kasatkina L.M., Lukovich V.V., Pilipenko V.V. Recognition of the person's voice by the classifier LIRA. *Control Systems and Computers*. 2006, N. 3, pp. 67–73.

94. Slipchenko S. V. Distributed representations for the processing of hierarchically structured numerical and symbolic information. *System Technologies*. 2005, N 6, pp. 134–141. (In Russian)
95. Rachkovskij D.A., Revunova E.G. Intelligent gamma-ray data processing for environmental monitoring. In: Intelligent Data Processing in Global Monitoring for Environment and Security. Kiev-Sofia: ITHEA, 2011, pp. 136–157.
96. Revunova E.G. Rachkovskij D.A. Training a linear neural network with a stable LSP solution for jamming cancellation. *Intern. Journal Information Theories and Applications*. 2005, Vol.12, N 3, pp. 224–230.
97. Revunova E.G., Rachkovskij D.A. Random projection and truncated SVD for estimating direction of arrival in antenna array. *Cybernetics and Computer Engineering*. 2018, N 3(193), pp. 5–26.
98. Hinton G. How to represent part-whole hierarchies in a neural network. *arXiv:2102.12627*. 2021, pp. 1–44.
99. Goyal A., Bengio Y. Inductive biases for deep learning of higher-level cognition. *arXiv:2011.15091*. 2020, pp. 1–42.
100. Greff K., van Steenkiste S., Schmidhuber J. On the binding problem in artificial neural networks. *arXiv:2012.05208*. 2020, pp. 1–75.

Received 17.03.2022

Рачковський Д.А.[1,2], д-р техн. наук, проф.,
головн. наук. співроб. відд. нейромережевих технологій
оброблення інформації, запрошений професор,
відділення компютерних наук, електротехніки та космічної техніки,
https://orcid.org/0000-0002-3414-5334, e-mail: dar@infrm.kiev.ua
Гриценко В.І.[1], член-кореспондент НАН України,
радник дирекції, https://orcid.org/0000-0002-6250-3987, e-mail: vig@irtc.org.ua
Волков О.Є.[1], канд. техн. наук, старший дослідник,
директор, https://orcid.org/ 0000-0002-5418-6723, e-mail: alexvolk@ukr.net
Гольцев О.Д.[1], канд. техн. наук,
в.о. зав. відд. нейромережевих технологій оброблення інформації,
https://orcid.org/0000-0002-2961-0908, e-mail: root@adg.kiev.ua
Ревунова О.Г.[1], д-р. техн. наук,
старш. наук. співр. відд. нейромережевих технологій оброблення інформації,
https://orcid.org/0000-0002-3053-7090, e-mail: egrevunova@gmail.com
Клейко Д.[3], доктор філософії (комп'ютерні науки),
наук. спів., https://orcid.org/0000-0002-6032-6155, e-mail: denis.kleyko@ri.se
Лукович В.В.[1],
наук. спів. відд. нейромережевих технологій оброблення інформації,
https://orcid.org/0000-0002-3848-4712, e-mail: vvl97@ukr.net
Осипов Є.[2], доктор філософії (комп'ютерні науки), професор,
відділення комп'ютерних наук, електротехніки та космічної техніки,
https://orcid.org/0000-0003-0069-640X, e-mail: evgeny.osipov@ltu.se
[1] Міжнародний науково-навчальний центр інформаційних
технологій та систем НАН України та МОН України,
40, пр. Академіка Глушкова, м. Київ, 03680, Україна
[2] Технологічний університет Лулео, 971 87 Лулео, Швеція
[3] Дослідні інститути Швеції RISE, 164 40 Кіста, Швеція

НЕЙРОМЕРЕЖЕВІ РОЗПОДІЛЕНІ ПОДАННЯ ДАНИХ ДЛЯ ШТУЧНОГО ІНТЕЛЕКТУ ТА МОДЕЛЮВАННЯ МИСЛЕННЯ

Вступ. Сучасний прогрес у галузі спеціалізованого штучного інтелекту пов'язано з використанням глибоких нейронних мереж. Однак вони мають ряд недоліків: потреба у величезних наборах даних для навчання, складність навчальних процедур, надмірна спеціалізація навчального набору, нестійкість до змагальних атак, відсутність інтеграції зі знаннями про світ, проблеми роботи зі структурами, відомі як проблема зв'язування або композиції. Подолання цих недоліків є необхідною умовою для просування від спеціалізованого штучного інтелекту до загального, що потребує розроблення альтернативних підходів.

Метою статті є огляд досліджень цього напряму, проведених у Міжнародному центрі протягом 25 років. Підхід до штучного інтелекту, що розробляється, випливає з ідей М.М. Амосова та його наукової школи. Також розглянуто зв’язки з напрямами гіпервекторних обчислень (HDC) та векторних символічних архітектур (VSA), а також з дослідженнями мозку.

Результати. Викладено концепцію розподіленого подання даних, включаючи HDC/VSA, які здатні подавати різні структури даних. Розглянуто розроблену парадигму асоціативно-проективних нейронних мереж: кодвекторне подання даних, операції суперпозиції та зв'язування, загальну архітектуру, перетворення даних різних типів у кодвектори, методи розв'язування задач та їхні застосування.

Висновок. Адекватне подання даних є одним з ключових питань штучного інтелекту. Основним напрямом дослідження, розглянутим у цій статті, є проблема подання різнорідних даних у системах штучного інтелекту в уніфікованому форматі на основі моделювання нейронної організації мозку та механізмів мислення. Розроблюваний підхід базується на гіпотезі розподіленого подання інформації в мозку та дає змогу подавати різні типи даних, від числових значень до графів, у вигляді векторів великої, але фіксованої розмірності.

Найважливішими перевагами розробленого підходу є можливість інтеграції та ефективного оброблення різних типів даних і знань, високий ступінь паралельності обчислень, надійність та стійкість до шумів, можливість апаратної реалізації з високою продуктивністю та енергоефективністю, оброблення даних на основі асоціативного пошуку за схожістю — подібно до того, як працює людська пам’ять. Це дає змогу уніфікувати методи, алгоритми та програмно-апаратні засоби для систем штучного інтелекту, підвищити їхню масштабованість за швидкістю та пам’яттю зі збільшенням обсягу та складності даних.

Дослідження створює основу для подолання недоліків сучасних підходів до створення спеціалізованого штучного інтелекту на основі глибоких нейронних мереж і відкриває шлях до створення загального штучного інтелекту.

Ключові слова: *розподілене подання даних, асоціативно-проективні нейронні мережі, кодвектори, гіпервекторні обчислення, векторно-символьні архітектури, штучний інтелект.*

DOI: https://doi.org/10.15407/kvt208.02.030

САВЧЕНКО-СИНЯКОВА Є.А, канд. техн. наук, с.н.с.,
с.н.с. відд. інформаційних технологій індуктивного моделювання,
https://orcid.org/0000-0003-4851-9664, e-mail: savchenko_e@meta.ua
Міжнародний науково-навчальний центр інформаційних технологій
та систем НАН України та МОН України,
пр. Акад. Глушкова, 40, м. Київ, 03187, Україна

ПОРІВНЯЛЬНИЙ АНАЛІЗ СТАТИСТИЧНОГО МОДЕЛЮВАННЯ ТА НАПРЯМІВ ШТУЧНОГО ІНТЕЛЕКТУ

***Вступ**. Сьогодні коло застосувань методів та інструментів машинного навчання розширюється дуже швидко, враховуючи загальну автоматизацію та цифровізацію. Використання методів та засобів машинного навчання для моделювання складних процесів дає змогу розв'язувати складні, практично нерозв'язні задачі. Проте інші методи математичного моделювання часто також досить успішно розв'язують задачу побудови моделі на основі вибірки експериментальних даних. Важливим завданням є визначення спільних та відмінних рис різних наукових напрямів штучного інтелекту для підвищення якості побудови математичних моделей.*

***Метою** статті є порівняльний аналіз засобів штучного інтелекту для побудови математичних моделей, таких як статистичне моделювання, машинного навчання, інтелектуальний аналіз даних та глибинне навчання, та визначення умов їхнього застосування.*

***Результати**. Проведено порівняльний аналіз машинного навчання та інших напрямів штучного інтелекту, який показав, що не дивлячись на схожість досліджуваних напрямів, вони мають низку суттєвих відмінностей. Ці напрями можна розрізнити за метою, можливістю перевірити або інтерпретувати отримані результати. На відміну від статистичного моделювання, застосування методів машинного навчання вимагає мінімальних людських зусиль, практично все навантаження лягає на програмне забезпечення.*

Проведене дослідження показало, що якщо ці два підходи можуть бути використані разом, це може привести до кращого результату, ніж кожен з підходів окремо. Результати порівняльного аналізу може бути використано для побудови системи підтримки прийняття рішень на основі індуктивного підходу та принципів метамоделювання.

***Висновки.** Виявлені спільні та відмінні риси засобів машинного навчання та статистичного моделювання визначають шляхи їхнього спільного використання для підвищення якості отриманого результату. Водночас треба враховувати, що за використання статистичного моделювання більше уваги приділяється поведінці моделі, а машинне навчання більше зосереджено на інтерпретуванні моделі та точності прогнозування.*

Ключові слова: *машинне навчання, статистичне моделювання, штучний інтелект, інтелектуальній аналіз даних, інтелектуальне моделювання, індуктивне моделювання.*

ВСТУП

Сьогодні, в час загальної комп'ютеризації та цифровізації, все більшу популярність здобувають методи та засоби математичного моделювання, за допомогою яких проводять аналіз великих обсягів даних, шукають в них взаємозв'язки та прогнозують показники на майбутнє. Одним з популярних засобів моделювання є методи машинного навчання, які дають змогу дуже швидко (часто в режимі онлайн) розв'язувати досить складні задачі.

Математичне моделювання залишається класичним інструментарієм отримання математичних моделей складних об'єктів та процесів. Але з появою сучасних продуктивних комп'ютерів з'явилась можливість обробляти величезні масиви даних, що раніше було неможливо. Це стимулювало розроблення нових підходів до аналізу такої інформації для побудови моделей за експериментальними даними. Наприклад, нейромережа навчається за вибіркою вхідних даних таким чином, щоб досягти високу точність обчислення значень вихідної змінної на основі вхідних даних.

Отже, напрям машинного навчання є підходом, що базується на математичній статистиці, але максимально використовує можливості сучасних комп'ютерів.

Метою цієї статті є аналіз спільних та відмінних рис різних наукових напрямів штучного інтелекту, таких як машинне навчання, статистичне моделювання та інтелектуальний аналіз даних, та визначення умов їхнього застосування.

ВИЗНАЧЕННЯ ПОНЯТЬ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ ТА МАШИННОГО НАВЧАННЯ

Моделювання в широкому сенсі визначають як метод дослідження процесів або явищ шляхом створення їхніх математичних моделей та подальшого дослідження цих моделей. *Метою класичного моделювання* є здобуття, оброблення, подання та використання інформації про об'єкти, які взаємодіють між собою та зовнішнім середовищем, тобто модель тут виступає як засіб пізнання властивостей й закономірностей поведінки об'єкта [1]. Моделювання включає етапи створення, дослідження та використання моделей об'єктів.

Методи моделювання широко використовують в різних сферах людської діяльності, особливо у проектуванні та керуванні, де основними є процеси ухвалення ефективних рішень на основі інформації про досліджувані об'єкт або процес. Основним призначенням моделі в завданнях керування є прогноз реакції об'єкта на різні керувальні впливи.

Розглянемо детальніше напрям **машинного навчання** (МН) як один із засобів моделювання процесів або об'єктів, та покажемо його місце серед напрямів штучного інтелекту. МН є підмножиною засобів штучного інтелекту, спрямованих на побудову моделей систем навчання за вибіркою експериментальних даних (навчальні дані) без використання явно запрограмованих етапів оброблення [2]. Алгоритми МН забезпечують побудову моделей, наприклад, для прогнозування [3], покладаючись на стандартні методи програмування, такі як об'єктно-орієнтоване програмування [4]. Водночас МН можна вважати спробою описати реальні процеси та системи шляхом апроксимації математичних функцій (рівнянь) [5].

Зазначимо, що основними властивостями методів МН є інтерпретація моделі, зменшення розмірності даних, можливість маніпулювання властивостями, перенесення властивостей на інші об'єкти та отримання нових знань тощо. До задач МН належать такі [5]: побудова регресії, класифікація, кластеризація, багатоваріантний пошук, оцінювання ймовірності, машинний переклад, виявлення аномалій тощо.

Формально завдання МН можна визначити так: комп'ютерна програма має видобути з досвіду *E* розв'язання деякого класу задач *T* та міри якості (продуктивності) *P* за умови, що якість *P* в *T* задачах покращується зі зростом досвіду *E* [6].

Але ж призначення МН набагато ширше, ніж тільки розроблення моделей для прогнозування та визначення коефіцієнтів моделі для опису певної предметної області [7]. МН зорієнтовано на вивчення експериментальних даних за допомогою комп'ютерних програм.

ПОРІВНЯННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ТА СТАТИСТИЧНОГО МОДЕЛЮВАННЯ

Статистичне моделювання (СМ) є підрозділом штучного інтелекту, який орієнтовано на пошук взаємозв'язків між змінними у вигляді математичних рівнянь, тобто СМ це – формалізація відношень між змінними за невеликого обсягу даних та кількістю атрибутів або спостережень, тому неможна нехтувати ймовірністю перенавчання моделей [4].

СМ працює за наявності припущень (обмежень), на відміну від алгоритмів МН, які в цілому позбавлені більшості таких припущень, зокрема для МН, як правило, не потрібно вказувати розподіл залежної або незалежної змінних. Тому методи МН можна застосовувати до різних типів даних, на відміну від методів СМ.

Схожість методів МН та СМ полягає у наявності схожих цілей, але вони не є однаковими напрямами. В [1] зазначено, що спільною рисою цих підходів є етап побудови моделі. Водночас у класичному моделюванні спочатку створюють модель, перевіряють її точність, визначають границі застосування і у разі необхідності здійснюють покращення результатів моделювання. МН дає змогу розв'язувати задачі завдяки застосуванню алгоритму навчання на великої кількості зразків, внаслідок чого можна досягти прийнятих результатів.

Основною відмінністю СМ та МН є те, що СМ заснована *на ймовірнісних просторах*. Статистичну теорію можна вивести з теорії множин [8], де числа можна згрупувати в підмножини, а потім розрахувати чи задати міру для кожної підмножини, щоб гарантувати, що сумарне значення цих мір буде дорівнювати одиниці.

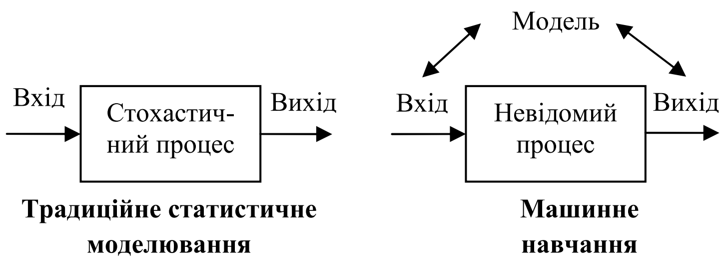


Рис. 1. Підходи до традиційного СМ та МН [9]

Друга значна відмінність СМ та МН полягає *в їхніх цілях*. Статистичні моделі використовують для пошуку та пояснення зв'язків між змінними, тоді як моделі МН створені для надання точних результатів на тестових даних. Хоча деякими статистичними моделями можна робити прогнози, точність цих моделей зазвичай не є найкращою, оскільки вони не можуть охопити складні зв'язки між даними. З іншого боку, моделі МН можуть надати кращі прогнози, але їх важче пояснити. МН є більш емпіричним підходом [9].

Отже, метою застосування СМ частіше є охарактеризувати взаємозв'язок між даними, а не робити прогнози, тоді як метою МН є отримати повторювані прогнози, і, в першу чергу, важливою є точність моделі, а її інтерпретування [9]. Інакше кажучи, МН дає інформацію про кінцеві результати, СМ — про взаємозв'язок між змінними, припускаючи, що дані є лінійним регресором з додаванням деякого випадкового шуму, який зазвичай має гаусовский характер [8].

Також відмінність цих методів вбачають в обсязі оброблюваних даних та у частці участі людини у побудові моделей [2]: МН добре працює з великими масивами даних, з великою кількістю змінних та спостережень; СМ зазвичай застосовують до менших масивів даних з меншою кількістю змінних.

Оскільки алгоритми МН можуть привести до перенавчання, то виникає необхідність поділу вибірки даних на навчальний та тестовий набір даних. Це не притаманно статистичним методам, оскільки у цьому випадку немає намагання мінімізувати емпіричний ризик.

Відмінність МН та СМ полягає в *способі перевірки моделей*. У традиційній культурі моделювання, побудована модель оцінюється за критерієм точності на тестових даних і це буде свідчити про точність прогнозу на нових даних. На відміну від цього, МН перевіряє модель на основі її попередньої точності, застосовуючи розподіл даних на дві нерівні частини в процесі побудови моделі: на одній частині здійснюється навчання за даними, на другій – перевірка, за якою визначають якість навченої моделі (налаштування на вхідні дані).

В основі методів МН покладено підхід «чорної скриньки», яка є невідомим випадковим процесом, порівняно з раніше передбачуваними простими випадковими процесами, які досліджуються за допомогою традиційного моделювання. За традиційним СМ весь сенс дослідження полягає в тому, щоб відкрити такі «чорні скриньки» для поглиблення знань про змодельовані природні процеси. , які лежать в їх основі. На рис. 1 схематично продемонстровано підходи традиційного СМ та МН.

У разі застосування СМ часто розглядають процес моделювання даних як шаблон для статистичного аналізу [10], розробляючи модель, що імітує процес, який намагаються формалізувати. Треба оцінити параметри і зробити висновки, але їх надають стосовно механізму роботи моделі, а не природного процесу, який вона імітує. МН базується безпосередньо на даних спостережень та не прив'язується до природного стохастичного процесу, тому може допомогти уникнути неправильних припущень щодо цього процесу [9].

Ще раз підкреслимо, що статистичні методи зосереджені на висновках про взаємозв'язки за досліджуваними даними, що досягається шляхом створення та налаштування конкретної ймовірнісної моделі, яка дає змогу обчислити кількісну міру впевненості в тому, що виявлений зв'язок описує «справжній» ефект, який навряд чи буде результатом шуму [11]. Крім того, якщо в наявності достатньо даних, ми можемо явно перевірити припущення і за потреби уточнити отриману модель.

Таблиця 1. **Порівняння базових понять СМ та МН [4]**

Риса	**Статистичне моделювання**	**Машинне навчання**
Очікуваний результат	Математична модель	Нейромережі, графи
Параметри для налаштування	Параметри	Ваги
Який механізм використовується	Налагодження	Навчання
Як працюють алгоритми	Продуктивність на тестовій вибірці	Узагальнення
Які завдання розв'язують	Оцінювання/Кластеризація	Навчання без учителя
На яких принципах основано	Математичні методи як основа, з припущеннями та обмеженнями	Вимагає менше припущень.
Спрямованість на використання	рівнянь	алгоритмів
Розмір вибірки даних	Невелика вибірка даних	Може працювати як з малими, так великими обсягами даних
Потреба в зусиллях людини	Зусилля людини потрібні	Мінімальне втручання людини
Властивості отриманої моделі	Найкраща «оцінка»	Гарні прогнозні властивості

Таблиця 2. **Порівняння особливих рис СМ та МН [13]**

Риса	Статистичне моделювання	Машинне навчання
Мета застосування	Формалізація взаємозв'язків між змінними у вигляді математичних рівнянь	Алгоритм, який може навчитися за даними, не покладаючись на програмування на основі правил.
На що спрямовано	Необхідно визначити функцію, яка найкращим чином опише вхідні дані перед налаштуванням цієї моделі (наприклад, лінійного, полінома тощо)	Не потрібно задавати функцію, оскільки алгоритми машинного навчання можуть автоматично вивчати складні шаблони даних
Точність прогнозування	Прогнозування результатів з точністю 85 %	Прогнозування результатів з точністю 85 %
Чи є тестування даних	Використання різних методів діагностування параметрів	Відсутність жодних статистичних тестів на значущість
Як відбувається тестування	Дані поділено на дві підвибірки: навчальну (70 %) та тестову (30%). Моделі будуються на даних навчальної вибірки, а тестуються як на загальній вибірці, так і для окремих змінних	Дані поділено на дві підвибірки: навчальну (50 %) та тестову (50%). Через відсутність діагностування алгоритмів МН, необхідно здійснювати навчання на двох вибірках даних (навчальній та тестовій), щоб забезпечити дві точки перевірки

Також СМ та МН відрізняються за *обчислювальною здатністю*. Класичне СМ було розроблено для даних із кількома десятками вхідних змінних і розмірами вибірки, які сьогодні вважалися б малими або помірними. Однак із збільшенням кількості вхідних змінних, модель, яка фіксує взаємозв'язки, стає складнішою.

Навпаки, МН зосереджується на прогнозуванні за допомогою алгоритмів навчання для пошуку закономірностей часто у великих та громіздких даних. Методи МН будуть найкорисніші, коли кількість об'єктів більша, ніж кількість вхідних змінних. МН робить мінімальні припущення щодо систем генерування даних; вони можуть бути ефективними, навіть якщо дані збираються без ретельно контрольованого плану експерименту та за наявності складних нелінійних взаємозв'язків у даних. Однак, незважаючи на переконливі результати прогнозування, є ймовірність ускладненої інтерпретацію отриманих результатів.

Навпаки, МН зосереджується на прогнозуванні за допомогою алгоритмів навчання для пошуку закономірностей часто у великих та громіздких даних. Методи МН будуть найкорисніші, коли кількість об'єктів більша, ніж кількість вхідних змінних. МН робить мінімальні припущення щодо систем генерування даних; вони можуть бути ефективними, навіть якщо

дані збираються без ретельно контрольованого плану експерименту та за наявності складних нелінійних взаємозв'язків у даних. Однак, незважаючи на переконливі результати прогнозування, є ймовірність ускладненої інтерпретацію отриманих результатів.

Треба враховувати відмінність у термінології у разі застосування СМ та МН [4].

Хоча є схожість між методичними підходами СМ та МН, але вона не завжди очевидна [12]. Обидва підходи вимагають мінімізації помилок, але використовують різні стратегії оптимізації для вдосконалення своїх алгоритмів. Кожний підхід має свої сильні сторони, і, як наслідок, можуть розглядатися як додаткова стратегія. Наприклад, МН може забезпечити певне розуміння процесу, який описують часові ряди, виділивши певний їхній клас, якому будуть відповідати різні методи побудови моделей. У табл. 2 показано риси, що притаманні і СМ, і МН, та їхні відмінності [13].

Отже, якщо потрібно вибрати алгоритм, який з високою точністю надаватиме прогнози, то слід вибрати МН. Якщо потрібно довести зв'язок між змінними або робити висновки за наявними даними, СМ буде найкращим [8].

ПОРІВНЯННЯ МАШИННОГО НАВЧАННЯ ТА ІНШИХ НАПРЯМІВ ШТУЧНОГО ІНТЕЛЕКТУ

До засобів штучного інтелекту також належить **інтелектуальний аналіз даних** (ІАД), який випередив МН майже на два десятиліття, спершу за назвою «виявлення знань у базах даних» (KDD) [14], а потім — Data Mining. *ІАД* — це сукупність методів виявлення за первинними даними раніше невідомих, нетривіальних, корисних на практиці та доступних для інтерпретації знань, які є необхідними для прийняття рішень в різних сферах людської діяльності [15]. ІАД використовують з 1930-х років, МН з'явився в 1950-х роках, вперше використаний в програмі гри в шашки.

ІАД є методом дослідження зібраних вхідних даних для визначення конкретного результату. Суть та мету технології ІАД можна охарактеризувати як технологію, яка призначена для пошуку в великих обсягах даних прихованих закономірностей. З іншого боку, у разі застосування МН здійснюється навчання для виконання складних завдань, з подальшим використанням зібраних даних для розв'язання все складніших завдань.

Отже, як МН, так і ІАД дають змогу відкривати закономірності в даних, тобто нові знання, але з різною метою: МН — здійснювати комп'ютерне навчання для розв'язання завдань; ІАД — для надання людині можливості досліджувати дані та інтерпретувати отримані результати. Водночас ІАД, в основному, зосереджується на розвідному аналізі за допомогою неконтрольованого навчання, тобто коли структура вхідних даних невідома. За [16], ІАД оперує змістовними завданнями, а МН — застосуванням математичної теорії, працює з алгоритмами [17], тому виникли такі терміни як, наприклад, «алгоритми машинного навчання в аналізі даних».

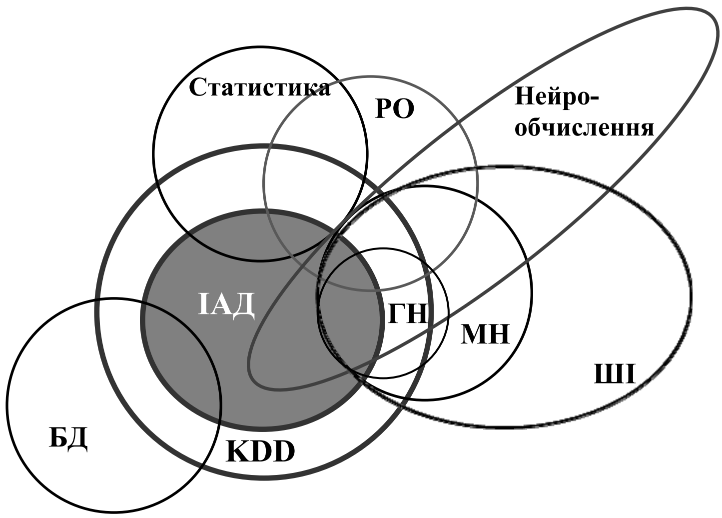


Рис. 2. Зіставлення машинного навчання та інших підходів до розв'язання завдань штучного інтелекту [9]

Отже, можна підсумувати особливості описаних напрямів.

Статистичне моделювання:

- більше, ніж ІАД, базується на статистичній теорії;
- більше зосереджується на перевірці гіпотез.

• *Машинне навчання:*

- має значно більшою мірою евристичний характер;
- концентрується на поліпшенні процесу навчання.

• *Інтелектуальний аналіз даних*

- характеризується інтеграцією теорії і евристик;
- дає змогу сконцентруватися на єдиному процесі аналізу даних, охоплює очищення даних, навчання, інтеграцію та візуалізацію результатів.

Data Sciences (DS) — наука про дані. Це обчислювальні та статистичні методи, які застосовуються до даних — як невеликих, так і великих наборів даних [8]. DS включає попереднє оброблення та аналіз даних, кодування, встановлення конвертора, з'єднання між БД, веб-сервісами тощо. Дослідницький або розвідувальний аналіз даних включає вивчення та візуалізацію даних, щоб допомогти фахівцю краще розуміти ці дані та робити з них висновки [18].

Ще однією групою методів МН є методи **глибинного навчання** або Deep Learning. Термін «глибинне навчання» (ГН) виник в середині 2000-х років, хоча окремі методи ГН відомі з 80-х років минулого сторіччя [20, 21]. Побудова нейронних мереж є найбільш широко використовуваним алгоритмом МН для прогнозування. Застосування нейронної мережі до часових рядів має різні варіанти залежно від структури та класу часових рядів і дає можливість обробляти складні структури даних [13].

Основою ГН є нейронна мережа з кількома рівнями для повторного навчання за вхідними даними [22]. Нейронна мережа імітує роботу людського мозку, що надає можливість вирішувати погано визначені проблеми. Прикладами засобів розв'язання таких завдань є програми розпізнавання зображень, мовлення та комп'ютерного зору.

Останнім часом все ширше використовують термін **інтелектуальне моделювання** (ІМ). Інтелектуальне моделювання визначено як процес побудови моделей об'єктів із застосуванням знань та інструментальних засобів, які забезпечують досягнення якості моделей на рівні кваліфікованого конструктора моделей (користувача, модельєра) [19]. Інтелектуальне моделювання включає набагато більше, ніж інструменти і методи розкриття шаблонів (залежностей) в даних. Модель має бути налаштована таким чином, щоб людина змогла зрозуміти та кількісно оцінити точність прогнозування для майбутніх (ще не отриманих) даних [7].

Для аналізу співвідношення МН, СМ та інших напрямів штучного інтелекту наведемо діаграму Венна [2], яка схематично демонструє відношення цих напрямів (рис. 2): ІАД — інтелектуальний аналіз даних, РО — розпізнавання образів, МН — машинне навчання, KDD (Knowledge Discovery in Database) — відкриття знань у базі даних, ГН — глибинне навчання. Ця діаграма візуально демонструє певні перетини напрямків штучного інтелекту між собою.

Kelleher, J. D. підкреслив, що першим розробником нейромережі ГН є О.Г. Івахненко, **метод групового урахування аргументів** (МГУА) якого вважається першою нейронною мережею, яка навчається за вхідними даними [21]. Перший загальний робочий алгоритм керованого навчання багаторівневої мережі персептронів був опублікований у [23]. У [24] було описано нейронну мережу з восьмима шарами, навченими методом групового урахування аргументів, який широко використовують і досі [25].

МГУА також може бути названо одним з алгоритмів інтелектуального моделювання, як метод синтезу моделей з автоматичним вибором структури та параметрів лінійних, нелінійних, різницевих та інших моделей на основі короткої вибірки даних в умовах невизначеності та неповноти вихідної інформації з метою виявлення невідомої закономірності функціонування досліджуваного об'єкта або процесу, інформація про яку неявно міститься у вибірці даних [19].

В алгоритмах МГУА поділ вибірки неявно (автоматично) забезпечує дотримання відомого принципу компромісу між складністю моделі та її точністю для побудови моделі оптимальної складності. Це дає можливість віднести МГУА до найефективніших методів обчислювального інтелекту, тому цей метод може бути основою для розроблення інструментальних засобів інтелектуального моделювання.

ПОЄДНАННЯ РІЗНИХ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ

Як вказувалось, метою застосування як МН так і СМ є навчання на основі даних. Але між цими підходами немає жорсткого розмежування. Деякі фахівці зі статистичного моделювання запозичили методи МН, рухаючись до об'єднаної області, яку вони називають «статистичним навчанням» (англ. statistical learning) [26]. Термін «статистичне навчання» є еквівалентом прогнозного моделювання [7].

Теорія статистичного навчання є основою для МН, яке базується на статистичній теорії та функціональному аналізі [27]. Теорія статистичного навчання має справу з проблемою статистичного висновку знаходження прогнозної функції на основі даних та успішно застосовується в таких галузях, як комп'ютерний зір, розпізнавання мови та біоінформатика.

Статистичне моделювання вимагає знати, як і чому була вибрана конкретна модель, а не як і чому робляться конкретні прогнози. Це не менш важливо, ніж інтерпретованість моделі.

Але висловлено думку [10], що область використання статистичного аналізу може зменшитись як через втрату актуальності, так і «крихкість» методів через ігнорування алгоритмічного підходу [7]. Тому спільне застосування таких підходів може давати кращі моделі, ніж кожен з підходів окремо [1].

Сьогодні застосування МН стало важливим елементом у прийнятті рішень. Це революціонізувало процес прийняття рішень завдяки мінімально можливому часу, необхідному для їх прийняття. Кожен рух людей, матеріалів, готової продукції тощо фіксується та зберігається як дані та використовується для прийняття рішень за допомогою засобів штучного інтелекту.

Поява та розвиток інструментів МН спростили статистичне прогнозування завдяки доступу до більшого обсягу даних порівняно з минулим.

Найкращим варіантом буде підхід, за яким враховуватиметься найкраще, що є в МН та СМ. Проведені дослідження демонструють, що значно корисніше використовувати ці підходи разом, ніж окремо, якщо є така можливість.

Отже, справжній фахівець з аналізу даних повинен мати в своєму арсеналі обидва підходи — МН та СМ. Сьогодні засоби МН не можна реалізувати без надійної статистичної бази [28]. Програмний код, написаний для полегшення роботи спеціаліста з аналізу даних, не скасовує потреби в глибокому розумінні теорії об'єктів чи процесів, що досліджуються.

ВИСНОВКИ

Проведений порівняльний аналіз різних напрямів штучного інтелекту, таких як машинне навчання, статистичне моделювання, інтелектуальний аналіз даних, глибинне навчання та інтелектуальне моделювання, показав, що не дивлячись на те, що всі ці напрями схожі між собою, вони мають значні відмінності та обмеження для опрацювання даних.

Показано, що головна відмінність цих підходів — їхня мета. Так методи СМ вимагають у розробника моделі розуміння взаємозв'язку змінних у рівнянні, щоб якнайкраще оцінити функцію виходу з прийнятною помилкою. Методи МН здійснюють навчання для розв'язання складних завдань, методи ІАД — надають людині можливість за допомогою комп'ютера розв'язувати складні задачі. Крім того, застосування методів МН вимагає мінімальних людських зусиль, практично все навантаження лягає на програмне забезпечення.

Проведений аналіз показав, що ті галузі досліджень, де застосовують спільно класичне (статистичне) моделювання і машинне навчання дають успішніші моделі ніж ті, які є результатом застосуванні цих підходів окремо. Результати проведеного дослідження може бути використано для побудови системи підтримки прийняття рішень на основі спільного використання статистичного моделювання і машинного навчання.

ЛІТЕРАТУРА

1. Поняття моделі та моделювання. Властивості та класифікація моделей. URL: https://web.posibnyky.vntu.edu.ua/fksa/2kvetnyj_komp'yuterne_modelyuvannya_system _procesiv/t1/11.htm (Дата звернення: 2 травня 2022).
2. Srivastava, T. Difference between Machine Learning & Statistical Modeling. URL: https://www.analyticsvidhya.com/blog/2015/07/difference-machine-learning-statistical-modeling (Дата звернення: 25 квітня 2022).
3. Sharma, D., & Kumar, N. A review on machine learning algorithms, tasks and applications. *International Journal of Advanced Research in Computer Engineering & Technology* (IJARCET). 2017. 6(10). P. 2278–1323.
4. Machine Learning vs Statistical Modeling. URL: https://www.youtube.com/watch?v=jaI5aheBOi0 (Дата звернення:2 травня 2022).
5. Most Common Machine Learning Tasks. URL: https://vitalflux.com/7-common-machine-learning-tasks-related-methods (Дата звернення: 25 травня 2022).
6. Mitchell, T. (1997). Machine Learning. New York: McGraw Hill. ISBN 0-07-042807-7. OCLC 36417892.
7. Difference Between Algorithm and Model in Machine Learning. URL: https://machinelearningmastery.com/difference-between-algorithm-and-model-in-machine-learning (Дата звернення: 12 квітня. 2022).
8. The Actual Difference Between Statistics and Machine Learning. URL: https://towardsdatascience.com/the-actual-difference-between-statistics-and-machine-learning-64b49f07ea3 (Дата звернення: 5 травня 2022).
9. Traditional Modeling vs. Machine Learning. URL: https://medium.com/thal%C4%93s/thales-traditional-modeling-vs-machine-learning-b7996342ea93 (Дата звернення: 12 квітня 2022).
10. Breiman, L. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 2001. 16(3). P. 199–231. doi: 10.1214/ss/1009213726.
11. Ij, H. Statistics versus machine learning. Nat Methods, 2018. 15(4). 233 p. https://www.nature.com/articles/nmeth.4642.
12. Dangeti, P. Statistics for machine learning. Packt Publishing Ltd., 2017. Packt Publishing. ISBN: 9781788295758.
13. Statistical Modeling vs Machine Learning. URL: https://blog.arkieva.com/statistical-modeling-vs-machine-learning/ (Дата звернення: 2 грудня 2021).
14. Mannila, H. Data mining: machine learning, statistics, and databases. Proceedings of 8th International Conference on Scientific and Statistical Data Base Management, 1996, pp. 2–9, doi: 10.1109/SSDM.1996.505910.
15. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine. 17(3). P. 37–37.
16. Van Calster, B., Verbakel, J. Y., Christodoulou, E., Steyerberg, E. W., & Collins, G. S. Statistics versus machine learning: definitions are interesting (but understanding, methodology, and reporting are more important). *Journal of clinical epidemiology*, 2019. 116.
17. Mitchell, Tom M. Machine learning and data mining. *Communications of the ACM*. 42.11, 1999. pp. 30–36.
18. Hripcsak, G., Duke, J. D., Shah, N. H., et all. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Studies in health technology and informatics*. 2015. 216. 574 p.
19. Goodfellow, I., Bengio, Y., & Courville, A. Deep learning. MIT press. 2016. 800 p. ISBN: 0262035618. DOI: https://doi.org/10.1007/s10710-017-9314-z.
20. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*. 61. pp. 85–117.

21. Степашко В.С. Концептуальные основы интеллектуального моделирования. *Управляющие системы и машины*. 2016. № 4. С. 3–15.
22. Kelleher, J. D. Deep learning. Cambridge, Massachusetts: The MIT Press; 2019. 280 p. DOI: https://doi.org/10.7551/mitpress/11171.001.0001.
23. Ивахненко, А.Г.; Лапа, В.Г. (1965). Кибернетические предсказывающие устройства. Киев: Наукова думка.
24. Ivakhnenko, A. (1971). Polynomial theory of complex systems. IEEE Transactions on Systems. *Man and Cybernetics* (4): 364–37.
25. Schmidhuber, J. My First Deep Learning System of 1991. Deep Learning Timeline 1960–2013.
26. Bzdok, D., Altman, N. & Krzywinski, M. Statistics versus machine learning. *Nature Methods*. 2018. 15. P. 233–234. https://doi.org/10.1038/nmeth.4642.
27. James, G., Witten, D., Hastie, T., & Tibshirani, R. An introduction to statistical learning. 2013. Vol. 112, New York: Springer. p. 18.
28. Statistical Modeling vs Machine Learning. URL: https://www.kdnuggets.com/2019/08/statistical-modeling-vs-machine-learning.html (Дата звернення: 2 квітня 2022).

Отримано 12.06.2022

REFERENCES

1. Ponyattya modeli ta modelyuvannya. Vlastyvosti ta klasyfikatsiya modeley. URL: <https://web.posibnyky.vntu.edu.ua/fksa/2kvetnyj_ komp'yuterne_modelyuvannya_system_procesiv/t1/11.htm> [Last accessed 2 May 2022] (In Ukrainian).
2. Srivastava, T. Difference between Machine Learning & Statistical Modeling. URL: <https://www.analyticsvidhya.com/blog/2015/07/difference-machine-learning-statistical-modeling/> [Last accessed 25 Apr. 2022].
3. Sharma, D., & Kumar, N. (2017). A review on machine learning algorithms, tasks and applications. *International Journal of Advanced Research in Computer Engineering & Technology* (IJARCET), 6(10), pp. 2278–1323.
4. Machine Learning vs Statistical Modeling. URL: <https://www.youtube.com/watch?v=jaI5aheBOi0> [Last accessed 2 Apr. 2022].
5. Most Common Machine Learning Tasks. URL: <https://vitalflux.com/7-common-machine-learning-tasks-related-methods> [Last accessed 25 May 2022].
6. Mitchell, T. (1997). Machine Learning. New York: McGraw Hill. ISBN 0-07-042807-7. OCLC 36417892.
7. Difference Between Algorithm and Model in Machine Learning. URL: <https://machinelearningmastery.com/difference-between-algorithm-and-model-in-machine-learning/> [Last accessed 2 Apr. 2022].
8. The Actual Difference Between Statistics and Machine Learning. URL: https://towardsdatascience.com/the-actual-difference-between-statistics-and-machine-learning-64b49f07ea3> [Last accessed 5 May 2022].
9. Traditional Modeling vs. Machine Learning. URL: <https://medium.com/thal%C4%93s/thales-traditional-modeling-vs-machine-learning-b7996342ea93> [Last accessed 12 Apr. 2022].
10. Breiman, L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199–231. doi: 10.1214/ss/1009213726.
11. Ij, H. (2018). Statistics versus machine learning. *Nat Methods*, 15(4), 233. https://www.nature.com/articles/nmeth.4642.
12. Dangeti, P. (2017). Statistics for machine learning. Packt Publishing Ltd.
13. Statistical Modeling vs Machine Learning. URL: <https://blog.arkieva.com/statistical-modeling-vs-machine-learning/> [Last accessed 2 Dec. 2021].
14. Mannila, H. (1996). Data mining: machine learning, statistics, and databases," Proceedings of 8th International Conference on Scientific and Statistical Data Base Management, pp. 2–9, doi: 10.1109/SSDM.1996.505910.

15. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17 (3), pp. 37–37.
16. Van Calster, B., Verbakel, J. Y., Christodoulou, E., Steyerberg E. W. & Collins G. S. (2019). Statistics versus machine learning: definitions are interesting (but understanding, methodology, and reporting are more important). *Journal of clinical epidemiology,* 116.
17. Mitchell, Tom M. (1999). Machine learning and data mining. *Communications of the ACM* 42.11, pp. 30–36.
18. Hripcsak, G., Duke, J. D., Shah, N. H., et all. (2015). Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. Studies in health technology and informatics, 216, 574 p.
19. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press. 800 p. ISBN: 0262035618. DOI:10.1007/s10710-017-9314-z.
20. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural networks, 61, pp. 85–117.
21. *Stepashko V.S.* (2016). Conceptual Fundamentals of Intelligent Modeling. *Upravlausie sistemy i masiny*. №4. pp. 3–15 (In Russian).
22. Kelleher, J. D. (2019). Deep learning. MIT press. DOI: https://doi.org/10.7551/mitpress/11171.001.0001.
23. Ivakhnenko, O.H.; Lapa, V.H. (1965). Kibernetychni prorochi prystroyi. Kyiv: Naukova dumka (In Russian).
24. Ivakhnenko, A. (1971). Polynomial theory of complex systems. IEEE Transactions on Systems, Man and Cybernetics (4): 364–37.
25. Schmidhuber, J. My First Deep Learning System of 1991. Deep Learning Timeline 1960–2013.
26. Bzdok, D., Altman, N. & Krzywinski, M., (2018). Statistics versus machine learning. Nature Methods. 15, pp. 233–234. https://doi.org/10.1038/nmeth.4642.
27. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: Springer.
28. Statistical Modeling vs Machine Learning. URL: <https://www.kdnuggets.com/2019/08/statistical-modeling-vs-machine-learning.html> [Last accessed 2 Apr. 2022].

Received 12.06.2022

Savchenko-Syniakova Ye.A., PhD (Engineering),
Senior Researcher of the Department for Information
Technologies of Inductive Modeling
https://orcid.org/0000-0003-4851-9664, e-mail**:** savchenko_e@meta.ua
International Research and Training Center
for Information Technologies and Systems
of the National Academy of Sciences of Ukraine
and Ministry of Education and Science of Ukraine,
40, Acad. Glushkov av., Kyiv, 03187, Ukraine

COMPARATIVE ANALYSIS STATISTICAL MODELING AND APPROACHES
OF ARTIFICIAL INTELLIGENCE

Introduction. Nowadays, the application of machine learning methods and tools is developing very rapidly, given the overall automation and digitalization. The use of machine learning methods and tools for modeling complex processes makes it possible to solve problems that were previously difficult or impossible to solve.

However other methods of mathematical modeling also make it possible to solve the problem of constructing a model based on a sample of experimental data. The task was to compare various scientific areas of artificial intelligence, such as machine learning, mathematical modeling, statistics, data mining and inductive modeling in terms of building mathematical models, to find out what common and distinctive features they have.

The purpose of the research is a comparative analysis of the areas of mathematical modeling, statistics and machine learning.

Results. A comparative analysis of machine learning and other approaches to solving artificial intelligence problems was carried out.

Conclusion. The conducted analysis shows that machine learning and mathematical (statistical) modeling are similar concepts, but not the same, and it does not depend on the purpose of applying algorithms based on these approaches. Where it is possible to use these two approaches together, this seems appropriate as it can increase the amount of data processed and increase the "understandability" of the resulting models.

Keywords*: machine learning, mathematical modeling, artificial intelligence, data mining, statistics.*

DOI: https://doi.org/10.15407/kvt208.02.044

BONDAR S.O., PhD student,
Researcher of the Intellectual Control Department.
https://orcid.org/0000-0003-4140-7985, e-mail: orangearrows@bigmir.net
SHEPETUKHA Yu.M., PhD (Engineering), Senior Researcher
Acting Head of the Intellectual Control Department
https://orcid.org/0000-0002-6256-5248, e-mail: shepetukha@irtc.org.ua
VOLOSHENYUK D.O., PhD (Engineering),
Senior Researcher of the Intellectual Control Department
https://orcid.org/0000-0003-3793-7801, E-mail: p-h-o-e-n-i-x@ukr.net
International Research and Training Center for Information Technologies
and Systems of the National Academy of Sciences of Ukraine
and Ministry of Education and Science of Ukraine.
40, Akad. Hlushkov ave., Kyiv, 03187, Ukraine.

USING OF HIGH-QUALITY POSITIONING TOOLS FOR HYBRID UNMANNED AERIAL VEHICLES AUTOMATIC CORRECTION UNDER THE LIMITED SPACE CONDITION

Introduction. *Original class of hybrid unmanned aerial vehicles is considered for multitask mission accomplishment at this article. Advantages of such vehicles usage for purposes that are always done by several different agents are considered. Perspective of the position precisioning for different tasks that could be done by unmanned aircrafts is analyzed.*

The purpose *of the paper is to universalize the process of surveillance, photo and video data collection and other missions that is provided by unmanned aerial vehicles today. The action of data precision during some periods of the misson accomplishment and increased specification for main targets of the mission could demonstrate brand new vector of the unmanned aerial vehicle usage and creation of the brand new domains for the unmanned aerial vehicles. Complex data gathering could help to avoid extra mediators and could simplify data processing on the next stages and also could do such data much more precise.*

Results. *The usable scenario of route for hybrid unmanned aerial vehicle and the model of it could be a proof of universal multitask unmanned aerial vehicle utilization. Such scenario unites several information missions of different scale and could provide data for several data centers that can use it for defferent problem solving just from one flight. Also it proves that utilization of such aircraft with an additional onboard precision block could be the next step at the mapping and digitalizing domains. Financial analysis of the market is provided for demonstration of the fact that such hybrid aircraft complex system would provide such scale as well as attention to the object details but be much cheaper then mapping and surveillance systems that are already existing.*

Conclusion. *A need for optimization of some problems that could be achieved by unmanned aerial vehicles leaded to the usage of hybrid vehicles that were represented at the paper. Complex design of such an aircraft could be a collateral disadvantage but the whole influence of the hybrid UAV usage for different tasks would optimize a lot more processes, devices and unnecessary equipment that would be needed for a large list of tasks at each domain UAVs are using right now from surveillance to agricultural tasks. Model of different scale purpose universal hybrid unmanned aerial system is a proof of the possibility to use just one single aircraft for a complex mission that needs different set of capabilities, features and equipment. Also such aircraft could provide much more certain results of missions and do it at lower pryce. Further developments could provide information about the most effective hybrid UAV type for such type of missions and provide game changing rules to the digitalizing and surveillance processes because of the new information gathering way.*

Keywords: *unmanned aerial vehicle, hybrid vehicle, positioning, multipurpose flight.*

INTRODUCTION

Throughout the unmanned aerial vehicles development history so as ordinary manned aircrafts, most attention has been paid to the development of aircrafts and helicopter (multi-propeller) types. The reliability and robustness of such designs played an important role in the construction of the first unmanned aerial vehicles. Also, limited range of tasks and lack of technological development sensibly eliminated the need to introduce other types of aircraft designs to unmanned aerial vehicles. However, in recent decades, the variety and focus of tasks that can be accomplished by unmanned aerial vehicles has expanded significantly, and there has been a need in different industries for aircraft that can provide combination of both basic aircraft types' benefits and universalize accomplishment of specialized tasks. It is exactly the case for combined aircraft types, as well as brand new aircraft types with experimental engines, to join missions of different orientation and accuracy tasks. All these new types of aircraft are combined under the concept of "hybrid". Hybrid unmanned aerial vehicles are becoming more and more popular in the world. However, today in Ukraine unmanned aerial vehicles of hybrid types are not so popular and have very narrow usage spectrum. But Ukrainian market has a high potential for usage of such UAVs for various civilian and military tasks.

A very promising area for the introduction of such unmanned aerial vehicles is the agrosphere, where aircraft of similar design could perform land irrigation more efficiently and solve other diverse problems, where the versatility of hybrid UAVs can be fully revealed. Geodesy and cartography can also be considered a promising field, where hybrid unmanned aerial vehicles can open up much more than conventional aircraft-type UAVs. In the military sphere, hybrid convertiplane aircrafts can be successfully used for border patrol tasks and pursuing violators.

PROBLEM STATEMENT

Recently, information collection by the unmanned aerial vehicles is a straight procedure that most of the time works as model one vehicle for one mission. But in some cases such type of work is not profitable. Studied area could have more than just one task to do and, also, such tasks could need different tools to accomplish the mission. Also, the mission could have several tasks that are needed diverse approach for its performance. So during the experiment the hybrid type of the unmanned aerial aircraft is needed to achieve a very special result of unity of the different scale tasks.

Different scale tasks could be taken at the surveillance and mapping domains – the field, where unmanned aerial aircrafts are having one of the vast usage parts already and could develop their perspectives even more.

The mission task would be in the simultaneous mapping of some large area part and the detailed data collection of some building that is situated at one point of such area. Task also includes further digitalizing procedure of gathered information and detailed 3D modelling of the studied area with the certain binding to the Earth coordinates.

The purpose of the paper universalize the process of surveillance, photo and video data collection and other mission that is provided by unmanned aerial vehicles today. The action of data precision during some periods of the misson accomplishment and increased specification for main targets of the mission could demonstrate brand new vector of the unmanned aerial vehicle usage and creation of the brand new domains for the unmanned aerial vehicles. Complex data gathering could help to avoid extra mediators and could simplify data processing on the next stages and also could do such data much more precise.

TYPES, ADVANTAGES AND DISADVANTAGES OF HYBRID UNMANNED AERIAL VEHICLES

Mostly whole hybrid aircraft concept could be classified by the working engine and aerodynamic plane position changing and different combinations of it, as could be seen on the Fig. 1:

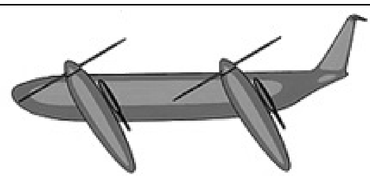
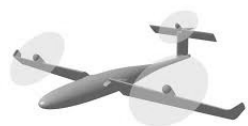
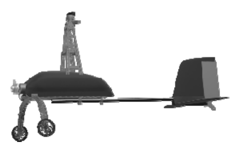
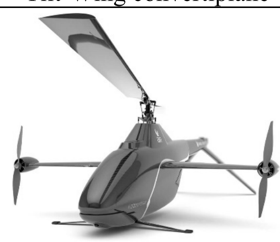
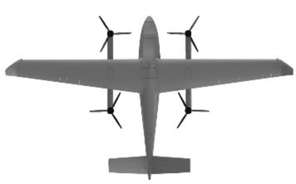
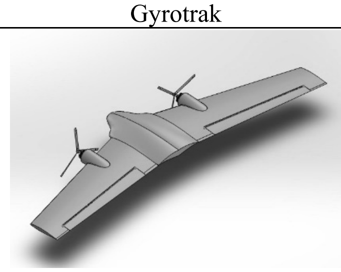
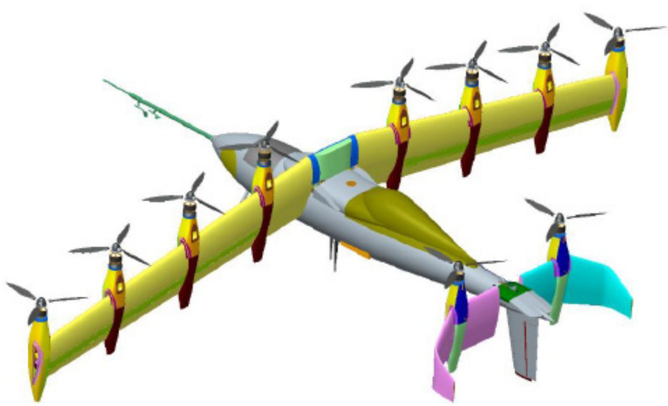
 Tilt-Prop convertiplane	 Tilt-Wing convertiplane
 Autogyro	 Gyrotrak
 Vertical Takeoff/Landing Aircraft with engine fixed in one plane	 Tailsitter

Fig. 1. The most popular hybrid aircraft types classification

Survey of hybrid unmanned aerial vehicles design features. According to the vast majority of market surveys, Tilt-Prop/Tilt-Wing convertiplanes are the most common type of a hybrid aerial vehicle in the industry. Basic hybrid unmanned aerial vehicle type that has been developed since the 80s of the twentieth century. The concept of such an aircraft was at first time used in the Bell prototype, "Eagle Eye" for the US Coast Guard purposes accomplishment [1]. Its concept has been developed over more than 10 years by leading aviation corporations and leaders of electronics and high technology domain. Convertiplanes have proven their effectiveness for video surveillance purposes, aerial photography and data collection for cartography, surveying, object searching and monitoring for that flight modes are needed to be changed.

There are vehicles of all sizes and diverse aerodynamic schemes. In particular: quadcopters with tilt propellers, quadcopters with tilt wings, aiplane-type aircrafts with the “whole tilting aerodynamic plane” (Fig. 2a).

a. Tilt-Wing convertiplane «Greased Lightning»



b. Vertical Takeoff and Landing Unmanned Aerial Vehicle prototype model

Fig 2. Unmanned convertiplane «Greased Lightning» as ultimate example of Tilt-Wing aircraft

Basic technology used at “Greased Lightning” aerodynamic plane (Fig. 2b) is an aerodynamic effectiveness that can be achieved in its “cruise” configuration. Electric engine on each prop satisfies the need for actuating shafts and conjunctions that are providing this aircraft configuration with distributed electric movement. Such design is developed for usage of the hybrid electric drive that includes small diesel engines, that are setting in motion alternating current generators to supply energy for the electric motor and to charge an onboard accumulator system.

Batteries provide power increase required for vertical takeoff and landing, as well as hanging. Many other new structural elements have been built in, such as hinged propellers to minimize drag when not in active use, so propulsion efficiency can be almost ideal in both “helicopter” and “airplane” flight conditions [2].

Vertical takeoff/landing aircrafts with engine fixed in one plane. The cheapest "compromise" type of hybrid aircraft is, in fact, a full-fledged hybrid of aircraft and helicopter types. It is an aircraft with propellers attached to the wings for vertical takeoff and landing. This design eliminates influence of the inertia moment on the rotary mechanism of the screw and completely solves the problem of limited reliability of this unit. However, such propeller arrangement affects the aerodynamic characteristics of the aircraft significantly and also increases fuel loss compared to similar “plane type” aircraft. The location of the engines also importantly affects device load capacity and its design rigidity. Such devices are used mostly for the delivery of small cargos, or for surveillance, aerial photography, or patrol purposes. Usually such aircrafts have small and medium size, when the impact of aerodynamics on the behavior of such structures is not so significant and can be offset by increased engine power and making minor changes to the power structure of the aircraft [3].

The prospect such aircraft usage to deliver larger loads, as well as the prospect of aerotaxi development, stimulated the design of such aircraft by integrating engines into the main powertrain design, such as placing engines inside the fuselage and wings to improve aerodynamic performance. Such integration is reflected in the “Trifecta” and “Accendance” unmanned aerial vehicles. Such schemes have a positive effect on the aerodynamic characteristics of the aircraft, can reduce fuel consumption and contribute to the creation of a more rigid structure, more stable behavior of the aircraft. Such improvements make it possible to transport larger loads and even people on board. The downside of such schemes is significant increase in the cost of production and maintenance of such devices, as well as increased impact of integrated engines vibration on the power structures that interact with them.

Also, the “Fixar” device could be added into the list that is built according to the original scheme with fixed screws mounted on inclined planes. In the future, such solution may be used in large devices.

In general, design of devices with fixed motors can differ significantly from each other. The versatility of the devices allows use of completely different schemes of plumage, wing placement, power structures etc.

Hybrid unmanned aerial vehicle of such scheme was suggested by the “Accendance” company for the aerotaxi purposes.

“Tailsitter”. A hybrid aircraft type that can take off and land vertically. Today it is the most promising type of unmanned aerial vehicle of hybrid type. Due to the smaller number of mechanical components compared to a

conventional convertiplane or autogyro, simpler engine design than vertical takeoff and landing vehicles, less weight of the entire structure, and less impact of moment of inertia when moving from one plane to another, this type of device has a wider range promising areas of implementation than hybrid designs of other types. Such devices have already established themselves as reliable means of delivery. Today, such aircraft are being tested to perform meteorological sounding tasks, as well as air taxis etc.

The tailsitter design is based on vertically placed aerodynamic elements and fixed engines, which are placed horizontally. Most of the existing aircraft and promising projects are designed according to the tailless and flying wing scheme and may vary according to the number of aerodynamic planes that can be installed in different quantity and in variable structure. Additionally, to the usual "tailless" and "flying wing" scheme, there are examples of bi- and triplanes, "three-pointed star", quadrofoil (quadroplane), inclined hexacryl (drone Amazon Prime).

Rotary motors placed on the nose of the structure, on its tail part, or on basic aerodynamic planes are used as the engine. Some designs (DelftAcopter [4]) have smaller auxiliary motors to adjust the position of the device on a plane perpendicular to the main.

“Autogyro”. The least popular and least stable type of hybrid unmanned aerial vehicles. Like its manned counterpart, unmanned autogyros remain a niche commodity and are now used to monitor objects and weather, and also as coast guard support devices and cargo delivery units. The type of autogyro for spraying chemicals in the fields can be singled out. The size of the aircraft varies from medium to large, but almost no small aircraft are used.

Autogyro designs are based on the presence of a propeller, the position and number of propellers may vary, as well as the number of aerodynamic planes and their position. For example, in addition to the standard concept of autogyros with a propeller and a push propeller between the usable volume compartment and the tail, or a propeller located at the front, two propeller concepts are developed as variant of the convertiplane scheme.

Recently, unmanned autogyros have been tested as couriers and assistants in the agricultural sector. Such tasks match the best to design features of autogyros. The possibility of autogyro development for use as an unmanned aerotaxi is also considered — an industry for which the design of autogyros and the possibility of safe landing, even with a failed engine, is the most suitable feature.

But the relative instability and design features of autogyros significantly limit their use. Compared to other types of hybrid aircrafts developed by dozens of aviation companies around the world, quantity of companies that develop autogyros is minor.

“Gyrotrak”. A separate type of hybrid unmanned aerial vehicle design based on the Gyrotrak platform that combines principles of autogyro and helicopter. The scheme is developed by Airial Robotics GMbH (Germany). The structure consists of bearing propeller and two pusher propellers located on the sides of the wing. In general, the model has a scheme that is more typical for helicopters, in particular, a pronounced tail with a V-shaped plumage at the end [5].

According to the designers, the autorotation of the propeller should provide energy savings and increased flight safety level, as well as range and autonomy of action, compared to multicopter systems. UAVs are also able to hang as helicopter does [6].

Hybrid UAV type has such benefits: does not require a catapult or runway for takeoff. It can be launched from the ground from relatively small areas, or hidden areas with difficult access; has a speed and range similar to airplane type aircraft; has the maneuverability of helicopter-type aircraft and has the ability to hang in the air; universal switching from one flight mode to another in a few seconds, which can allow such aircraft to perform several tasks simultaneously; flexibility of work; does not require a full-fledged chassis.

Along with the obvious advantages of hybrid UAVs also have a number of disadvantages: more complex design, compared to aircraft and helicopter type; lower reliability than airplane and helicopter types due to suppression of the inertia moment during the transition of engines from mode to mode; less time in the air than in aircraft with similar dimensions and characteristics; higher production cost than airplane and helicopter types; higher maintenance price of hybrid devices.

As a result of the last two points, it is risky to use such aircraft for tasks with an increased risk of device damage, or its loss. To reduce risk factors, it is necessary to ensure effective adjustment of UAVs on time. The importance of this adjustment is increasing due to the fact that the long-term use of hybrid unmanned aerial vehicles involves their widespread use for emergencies, accidents at infrastructure, data collection for mapping and photo and video data collection purposes. It is for such tasks that the automatic adjustment of the position of the unmanned aerial vehicle in space is especially important.

Hybrid unmanned aerial vehicles at limited space. In many cases, a significant additional factor to consider is the limited space in which UAVs move. Adjusting the position in confined spaces is necessary for unmanned aerial vehicles to perform procedures and tasks that require extreme precision due to the importance of the task or the risk of damage to the aircraft.

Under the definition of limited space specified area limited with physical obstacle, energetical or optical defense devices or limited by the special program tool in the aircraft that accomplishes some mission software is understood. Aircraft design factors may also limit the possible space: the range of the signal controlling the flight of the unmanned aerial vehicle and the amount of fuel that the aircraft can take on board at departure.

Thus, the concept of physically limited space includes the terrain and objects in the flight path of an unmanned aerial vehicle. A promising method of studying physically confined space is computer modeling of its most important components. During the modelling process, it is necessary to consider several types of physical space constraints to demonstrate the versatility of the use of unmanned aerial vehicles and possibility of its operation during the different scale task performance.

HYBRID UNMANNED AERIAL VEHICLE USAGE ALGORYTHM

The usage of universal hybrid type UAVs makes it possible to respond as quickly as possible to unforeseen events and perform high-precision tasks over long distances. In addition to responding to emergencies, such UAVs can be used to refine the positioning of objects at a great distance from the operator. A promising area of application is also operations for which it is important to cover a large area and identify objects in the area that require high accuracy of location, and description of specific design features.

But for the successful solution of all these problems it is necessary, first of all, to ensure high-precision positioning of the hybrid UAV.

Summarizing the information provided in the specific literature, we can propose the following structure of the hybrid UAVs usage to solve following problems:

- definition of an approximate object location;
- UAV takeoff to the destination point in the “airplane” mode;
- arriving to the destination point and object localization in general;
- transition to the helicopter mode;
- detailed positioning data compilation;
- receiving information about the status and specification of the required object construction with enhanced level of the information certainty;
- an opposite transition to the “aircraft” mode;
- returning to the initial base point.

The proposed structure was used in the test implementation of the hybrid UAV exploitation scenario for the certain part of the area mapping task performing.

HYBRID UNMANNED AERIAL VEHICLE TEST SCENARIO

The small VTOL Freeman 2300 aircraft, one of the most affordable hybrid unmanned aerial vehicles in the world, was selected as a test vehicle example [7]. The main characteristics of this aircraft are that it is built according to the vertical takeoff and landing scheme, it is adapted for the use of onboard cameras, it can carry up to 1.5 kg payload, which is sufficient for the installation of professional equipment for photo and video data collection, mapping and positioning mission accomplishment. Such aircrafts can be situated on limited landing zone. Starting complex, special tools or the runway availability nonobligatoriness for such unmanned aerial vehicle is a significant benefit. Aircraft starting procedure can be done without additional staff interference, also, proximate operator eye control during the whole start procedure is not necessary. Operator can be located elsewhere and monitor the behavior of UAVs by on-board video cameras and sensors that are also on board the device.

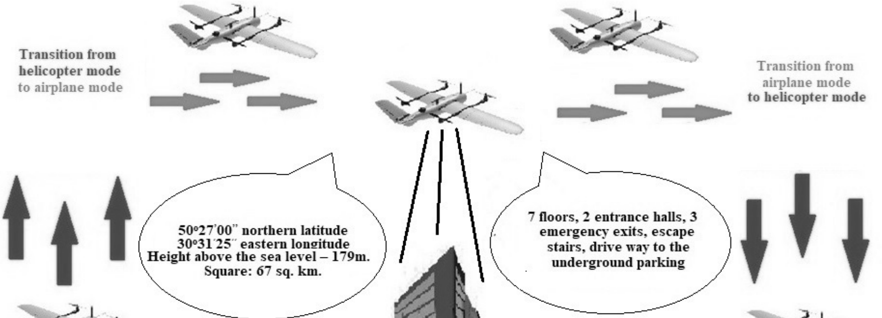


Fig. 3. The scheme of an approximate flight scenario of a hybrid unmanned aerial vehicle for future construction of a house 3D model and its exact positioning on the map

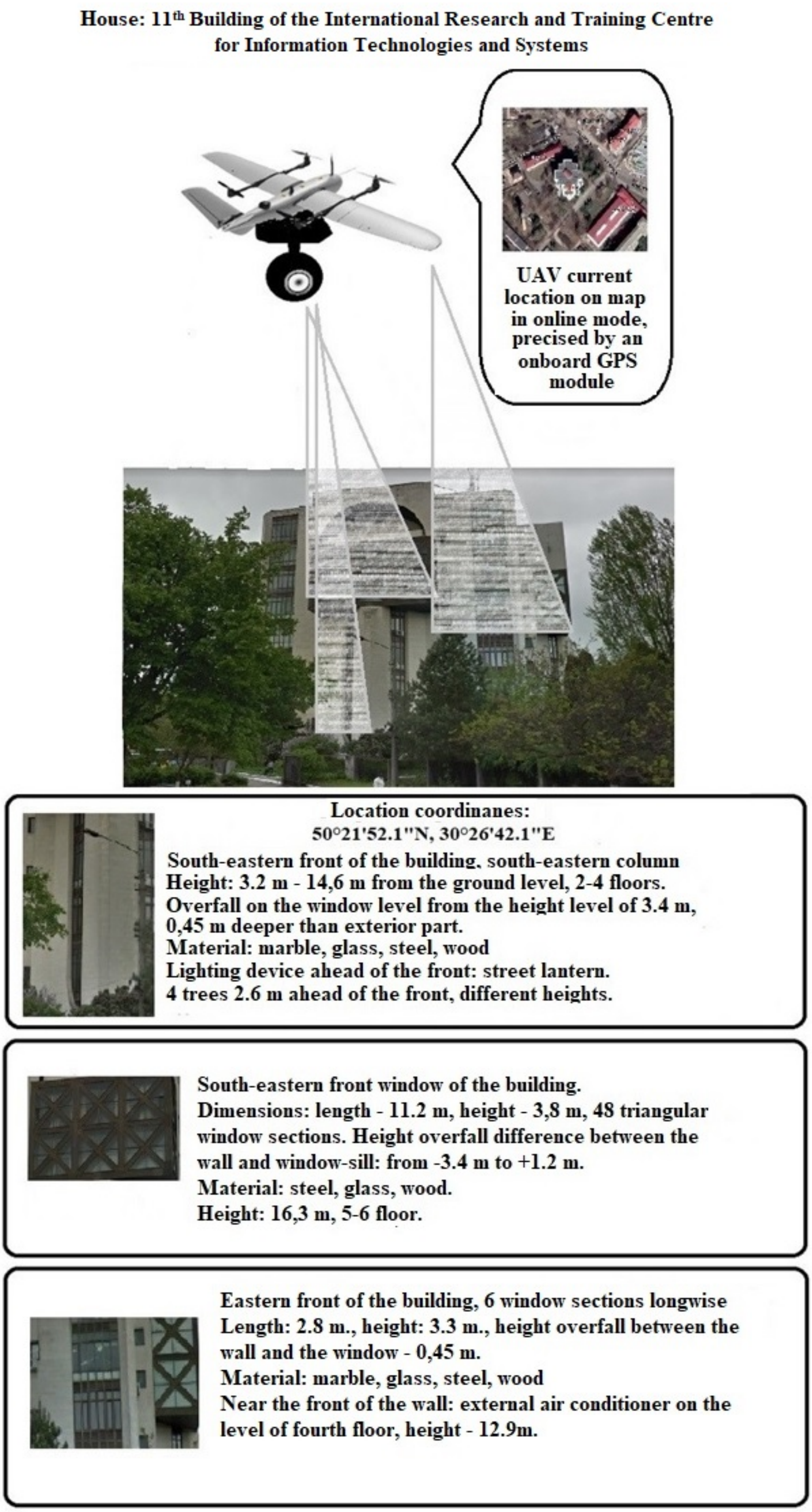


Fig 4. Unmanned Aerial Vehicle information gathering for the purpose of following house 3D model construction

During the tests a UAV flight with a mapping misson at the point 50 km from the aircraft start point has been simulated (Fig 3). Construction of the landscape map includes detailed reference to coordinates with inaccuracy value not more than 1m (that is certain enough correspondingly to the available landscape maps) and mapped landscape area 3D model construction.

The aircraft was operated from a control station located near the take-off point. Departure took place vertically from the base and at an altitude point around 800 m above the ground the aircraft switched to horizontal “airplane” flight mode. The transition from one mode to another took 8-10 seconds, which allows you to do it with an accuracy of 1-2 meters in height. The altitude stabilization mode made it possible to return to the exact height of 800 m at the end of the transition maneuver. With the help of additional helicopter propellers that action does not affect the horizontal flight, the position was adjusted to the course without changing the plane of flight. The software adjustment was carried out by obtaining updated information from the autopilot, which aims to optimize the flight path of the UAV and is based on the method of managing the full unmanned aerial vehicle energy. The use of software and hardware complex of the control system of unmanned aerial vehicles has detailed and optimized the redistribution of the aircraft energy [8]. It also made possible to reduce the error of information perception during the collection data on photo and video equipment installed on board the aircraft. When the aircraft approached the measurement site, the hybrid UAV again switched to "helicopter" response mode and "hung" in the air to clarify the positioning of objects and the device itself during the necessary data receiving process. An example of data collection to clarify the particular object positioning for further 3D model construction and clarifying its location is represented on Fig. 4.

ANALYSIS OF TEST RESULTS

One of the important advantages of hybrid UAVs is that in presence of terrain differences, such a device can measure with high accuracy the difference in altitude and the type of obstacle in the vision area of the onboard cameras. If such measurements are required, the overflight device can switch to helicopter mode and study the required object in detail at close range. Detailed flight of the selected object and data collection from all possible points allows building a high accuracy model of any selected object. Additionally, emergency sensors help to avoid contact with obstacles and to protect the aircraft from unwanted damage.

A critical advantage of hybrid aircraft usage is that high-precision positioning makes it possible to operate such an unmanned aerial vehicle in fully autonomous mode, and the presence of an autopilot with virtual control of course, pitch and yaw in both modes significantly reduces deviation from the specified route [9]. During the work in autonomous mode, the aircraft can perform the task of collecting photo and video information independently on a pre-programmed route (Fig. 5). In case of obstacle appearance that has not been previously described during the construction of the route, the aircraft is able to build the adjusted route and enter information about its change in the system. Thanks to a combination of both aircraft and helicopter capabilities, a hybrid unmanned aerial vehicle can also perform a detailed study of the objects assigned to the study.

At the same time, upon completion of a study, a hybrid unmanned aerial vehicle can return to general site monitoring till the next target object appears, which requires a detailed study and construction of an accurate virtual model based on the collected data. At the same time, the refinement module and the system for adjusting the position of the unmanned aerial vehicle in space will provide it with the most optimal flight trajectory between objects, as well as reduce energy consumption compared to other examples of unmanned aerial vehicles.

From the aforesaid, it can be concluded that the presence of such a combination as a high-precision positioning device and a hybrid unmanned aerial vehicle makes it possible to study in detail the hard-to-reach places on the Earth's surface. Such as, for example, the interiors of houses, deepening quarries, caves and mines, large-diameter pipelines, reservoirs, tunnels and power grids, forests and agricultural land.

Thus, such unmanned aerial vehicles with an installed unit of positioning precision and adjusting the position of the aircraft in space are indispensable for mapping, data collection from the state of individual buildings, accounting for architectural management, creating models of functioning and development of cities, transport models, control the state of infrastructure facilities. It is also possible to perform tasks in response to emergencies; accidents, fires, evacuations from closed premises. It is also important to make a detailed assessment of losses and ways to reduce them when heavy equipment cannot access the required areas. It will also be easier to comprehensively assess the factors that have led to such emergencies and ways to prevent them.

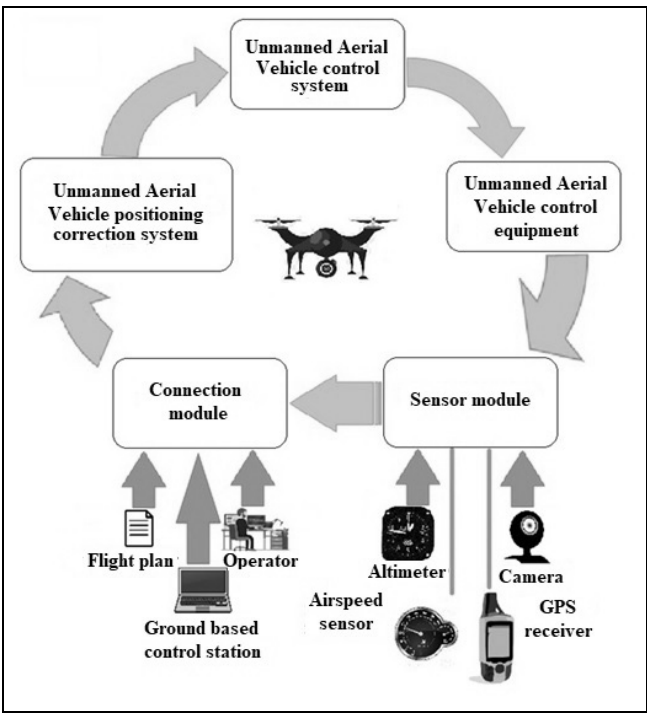


Fig. 5. Scheme of Unmanned Aerial Vehicle with positioning in space correction system working process.

USAGE PERSPECTIVES

Thus, it can be concluded that usage of unmanned aerial vehicles for such tasks is quite justified at this stage of development of observation methods and technologies. Satellites, helicopter and aircraft UAVs, and ground equipment can be used for the described positioning scenarios. Each of these technical measures is characterized by its specific scope, has its advantages and disadvantages.

Thus, satellites are used for general mapping of the area with relatively low accuracy and little attention to detail and conditional accuracy of coordinates. The quality of the satellite image is insufficient without external intervention and clarification by above-ground photography. Therefore, maps obtained by the help of satellites can only be used as a starting point for data collection by unmanned aerial vehicles. Additionally, satellite systems can serve as consumers of updates collected by unmanned aerial vehicles. The disadvantage of satellite technology is that such technology is complex, and as a result, too expensive to be fully used for mapping and clarifying locations and objects, creating models.

Another technological solution involves data collection by unmanned aerial vehicles of helicopter and airplane types. Such solution is very popular now, but lacks the versatility that hybrid drones can offer, including vertical takeoff and landing vehicles and “tailsitters”. These tools can offer increased required data collection accuracy about points and obtaining more information about the same area. The versatility of this concept of obtaining information about the area allows you to reduce the number of aircraft required to collect data from a particular area. It also reduces the number of departures required to maintain accuracy that is several times higher than what modern systems can offer today. Additionally, it is possible to reduce area size required for deployment of the complex, storage of equipment and take-off and landing procedures accomplishment.

Proposed approach usage advisability techno-economical feasibility. The proposed approach is not only determined by scientific novelty, has technological advantages, but also is justified economically.

According to available data, one Freeman 2300 VTOL aircraft with a set of on-board cameras and a suspended camera with 360-degree image coverage is needed to deploy one mapping system that could cover up to 160 km around the landing zone. The range of device prices can vary from $ 1,000 to $ 15,000 depending on the type of onboard equipment. Extended control station for the device, which can cost up to $ 7,500, GPS-tracker for $ 150–350. Totally, the cost of a complex consisted of one device for object monitoring located within a radius of 80 km is $ 23,000 [10]. Such amount is significantly less than the usual price of the equipment used for surveillance purposes at present. For example, the cost of aircraft photography and mapping system with less versatility of the system are often several times higher. And the cost of military drone implementation on the market is 10–12 times higher. If we also taking into account price of participation in global space programs, such numbers would differ from the complex presented here by thousands of times.

Along with a significant price reduction, proposed approach will ensure high-precision system operation. Modern equipment usually involves usage of space technologies and ground-based refinement systems. Due to the GPS-tracker installation and use of high-quality photographic equipment, positioning objects accuracy on the map can be increased by 1 to 10%, depending on the type of object, which is a significant clarification compared to the techniques used to solve most cartographic tasks today. Additionally, technology will be perfect enough to obtain accurate data to create a 3D model of objects. All this will be a significant step towards creation of accurate 3D models of cities, which in turn will improve the prevention of emergencies, facilitate tasks of general planning, development of certain areas, functioning of sewerage systems and public transport etc.

CONCLUSION

Hybrid unmanned aerial vehicles are the most promising tools for positioning, mapping, monitoring and modeling. The design of such aircraft makes it easier to maneuver in space and explore hard-to-reach objects.

The position adjustment system, especially in limited spaces, can ensure the safety of such an aircraft both at high speeds when operating in "airplane" mode and during the "helicopter" flight in dangerous proximity to the object that needed to be investigated.

Adjusting the position of the UAV in space is possible only by accurately calculating its location and using the positioning refinement module. This increases not only the accuracy of location calculation, but also the accuracy of observated object coordinates determination. Main domains of this complex improvement are ensuring high accuracy of data collection and positioning of the object under investigation, flexibility in the use of different types of data and the ability to use the complex to perform tasks of different scales simultaneously.

The paper contains an example of a hybrid aircraft flight scenario and shows the possibility of its application for tasks requiring different speed modes, different ranges of observation tasks, data collection simultaneously for several purposes that can be located at considerable distances from each other and perform general observation tasks by location area paralelly.

It proves that usage of modern hybrid aircrafts can be completely autonomous, safe, and have quite large working time periods. It may allow you to collect detailed information and cover a larger area / perform fewer flights than during the usage of other UAV types. Also, the use of such systems will optimize the performance of tasks of different directions, reduce the number of different unmanned aerial vehicles to perform different types of tasks and, at the same time, increase the accuracy of any type of tasks related to positioning, surveillance, delivery etc.

REFERENCES

1. FAS Intelligence Resource Program. Eagle Eye UAV. URL: https://irp.fas.org/program/collect/eagle-eye.htm
2. Flores Reyes A., Flores Colunga G.R. Design and Development of an UAV with Hybrid Flight Capabilities. Leon, Guanajuato, Mexico, 2018. URL: https://cio.repositorioinstitucional.mx/jspui/bitstream/1002/800/1/17444.pdf
3. Saeed A.S., BaniYounes A., Cai Ch., Cai G. A Survey of Hybrid Unmanned Aerial Vehicles. *Progress in Aerospace Sciences Journal*. Oxford, United Kingdom, 2018. Vol. 98, pp. 98–105. https://doi.org/10.1016/j.paerosci.2018.03.007
4. delftAcopter: innovative single-propeller hybrid drone URL: https://tudelftrobotics-institute.nl/robots/delftacopter

5. Douglas A. Airial Robotics develops "new type" of UAV for global commercial drone market. Commercial Drone Professional website. 2020. URL: https://www.commercialdroneprofessional.com/breaking-news-airial-robotics-develops-new-type-of-uav-for-global-commercial-drone-market/
6. Barskiy R. Gyrotrak: brand new hybrid UAV concept. *Science and technic*. 2020. URL: https://naukatehnika.com/gyrotrak-novaya-gibridnaya-koncepcii-bespilotnikov.html
7. Makeflyeasy Freeman 2300 Tilt VTOL Aerial Survey Carrier Span Fpv Rc Fix-wing Model drone Wing 2300mm UAV mapping Long range pryce. URL: https://www.uavmodel.com/products/makeflyeasy-freeman-2300-tilt-vtol-aerial-survey-carrier-span-fpv-rc-fix-wing-model-drone-wing-2300mm-uav-mapping-long-range
8. Grytsenko V.I., Volkov O.Ye., Komar.M.M. et al. Modern unmanned aerial vehicle automatic control systems intellectualization. *Cybernetics and Computer Engineering,* 2018, № 1(191), pp. 45–59. URL: http://kvt-journal.org.ua/834/ (in Ukrainian)
9. Volkov O.Ye., Grytsenko V.I., Komar M.M. et al. Integral Adaptive Autopilot for an Unmanned Aerial Vehicle. *AVIATION: Scientific journal: scientific article*. Vilnius, Lithuania, 2018. Vol. No 22. pp. 129–195.
10. Makeflyeasy Freeman 2300 Specification & Options URL: https://aliexpress.com/item/10000223137957.html

Received 06.04.2022

ЛІТЕРАТУРА

1. FAS Intelligence Resource Program. Eagle Eye UAV. URL: https://irp.fas.org/program/collect/eagle-eye.htm
2. Flores Reyes A., Flores Colunga G.R. Design And Development Of An Uav With Hybrid Flight Capabilities. Leon, Guanajuato, Mexico, 2018. URL: https://cio.repositorioinstitucional.mx/jspui/bitstream/1002/800/1/17444.pdf
3. Saeed A.S., BaniYounes A., Cai Ch., Cai G. A survey of hybrid Unmanned Aerial Vehicles. *Progress in Aerospace Sciences Journal*. Oxford, United Kingdom. 2018, Vol.98, pp. 98–105. https://doi.org/10.1016/j.paerosci.2018.03.007
4. delftAcopter: innovative single-propeller hybrid drone URL: https://tudelftroboticsinstitute.nl/robots/delftacopter
5. Douglas A.Airial Robotics develops "new type" of UAV for global commercial drone market. Commercial Drone Professional website. 2020. URL: https://www.commercialdroneprofessional.com/breaking-news-airial-robotics-develops-new-type-of-uav-for-global-commercial-drone-market/
6. Р. Барський Gyrotrak: новая гибридная концепции беспилотников. *Наука и техника,* 2020. URL: https://naukatehnika.com/gyrotrak-novaya-gibridnaya-koncepcii-bespilotnikov.html
7. Makeflyeasy Freeman 2300 Tilt VTOL Aerial Survey Carrier Span Fpv Rc Fix-wing Model drone Wing 2300 mm UAV mapping Long range pryce. URL: https://www.uavmodel.com/products/makeflyeasy-freeman-2300-tilt-vtol-aerial-survey-carrier-span-fpv-rc-fix-wing-model-drone-wing-2300mm-uav-mapping-long-range
8. В.І. Гриценко, О.Є. Волков, М.М. Комар, Ю.П. Богачук. Інтелектуалізація сучасних систем автоматичного керування безпілотними літальними апаратами. *Кибернетика и вычислительная техника*, 2018. № 1(191). С. 45–59. URL: http:// http://kvt-journal.org.ua/834/
9. Волков О.Є., Гриценко В.І., Комар М.М., Волошенюк Д.О. Integral Adaptive Autopilot for an Unmanned Aerial Vehicle. *AVIATION: Scientific journal: scientific article.* Vilnius, Lithuania, 2018. Vol. No 22. С. 129–195.
10. Makeflyeasy Freeman 2300 Specification & Options. URL: https://aliexpress.com/item/10000223137957.html

Отримано 06.04.2022

Бондар С.О., аспірант,
наук. співроб. відд. інтелектуального управління
ORCID: 0000-0003-4140-7985, e-mail: orangearrows@bigmir.net
Шепетуха Ю.М., канд. техн. наук, старш. наук. співроб.,
в.о. зав. відд. інтелектуального управління
ORCID: 0000-0002-6256-5248, e-mail: shepetukha@irtc.org.ua
Волошенюк Д.О., канд. техн. наук,
старш. наук. співроб. відд. інтелектуального управління
ORCID: 0000-0003-3793-7801, E-mail: p-h-o-e-n-i-x@ukr.net
Міжнародний науково-навчальний центр інформаційних
технологій та систем НАН України і МОН України,
40, пр. Академіка Глушкова, м. Київ, 03187, Україна

ВИКОРИСТАННЯ ВИСОКОЯКІСНИХ ІНСТРУМЕНТІВ ПОЗИЦІОНУВАННЯ ДЛЯ АВТОМАТИЧНОЇ КОРЕКЦІЇ ГІБРИДНОГО БЕЗПІЛОТНОГО ЛІТАЛЬНОГО АПАРАТУ В УМОВАХ ОБМЕЖЕНОГО ПРОСТОРУ

Вступ. Для виконання завдань з різною кількістю різномасштабних підзавдань ефективно використовують клас так званих гібридних безпілотних літальних апаратів (БпЛА). Безпілотні літальні апарати такого класу мають низку переваг, зокрема використання таких транспортних засобів для цілей, що завжди виконуються кількома різними виконавцями, які часто навіть не зв'язані між собою. Проводиться спроба виконати аналіз перспективи уточнення місцезнаходження досліджуваного об'єкту для різних завдань, які можуть виконуватися безпілотними літальними апаратами.

Метою статті є універсалізація процесу спостереження, збору фото- та відеоданих та інших завдань, які сьогодні забезпечують безпілотні літальні апарати. Процес визначення точності даних протягом окремих періодів виконання місії та підвищення специфікації основних цілей місії може започаткувати абсолютно нову сферу використання безпілотних літальних апаратів та підштовхнути розвиток абсолютно нових напрямів застосування безпілотних літальних апаратів. Комплексний збір даних виключає наявність додаткових посередників і міг би спростити процедуру оброблення даних на наступних етапах операції, а також забезпечувати набагато точнішою інформацією потенційних споживачів даних картографування, геолокації, диспетчеризації та універсуалізувати джерело отримання інформації для них.

Результати. Сценарій використання саме гібридного безпілотного літального апарату та його реалізація в 3D середовищі можуть слугувати гарним прикладом виконання універсальної багатоланкової задачі з різним масштабом дистанцій для кожної окремої ланки. Такий сценарій об'єднує кілька інформаційних місій різного масштабу і може надати дані для кількох центрів оброблення даних, які можуть використовувати їх для розв'язання різних завдань лише під час одного польоту. Також це доводить, що використання такого апарату з додатковим бортовим уточнювальним блоком може стати наступним витком розвитку галузей оцифровування даних та картографії. Фінансовий аналіз ринку надається для демонстрації того, що така гібридна авіаційна комплексна система забезпечить виконання завдань різного масштабу, точніше опрацює деталі об'єкту-цілі і водночас буде значно дешевшою за наявні системи картографування та спостереження.

Висновок. Потреба в оптимізації деяких завдань, які могли б виконувати безпілотні літальні апарати, привела до впровадження гібридних транспортних засобів, наданих/описаних/проаналізованих в роботі. Складна конструкція такого літального апарату може бути побічним недоліком, але вплив використання гібридного БпЛА для виконання різних завдань оптимізує набагато більше процесів, скоротить витрати на побічні пристрої і обладнання, що потрібно для великого переліку завдань у кожній галузі, де використовуються БпЛА сьогодні, від спостереження та аерофотозйомки до сільськогосподарських та військових завдань. Модель універсальної гібридної безпілотної літальної системи різного масштабу є доказом можливості використання лише одного літального апарату для виконання складної місії, яка потребує різного набору можливостей, функцій та обладнання. Також такі літальні апарати могли б забезпечити набагато точніші і місткіші дані за результатами виконання місій за рахунок менших матеріальних витрат. Подальші розроблення допоможуть отримати інформацію про найефективніший тип гібридного БпЛА для місій такого типу і сформулювати абсолютно нові постулати в галузях процесів цифровізації та спостереження, використовуючи новий спосіб збору інформації.

Ключові слова: *безпілотний літальний апарат, гібридний апарат, позиціонування, багатоцільовий політ.*

Medical
and Biological Cybernetics

DOI: https://doi.org/10.15407/kvt208.02.060

ARALOVA N.I.[1], DSc (Engineering), Senior Researcher,
Senior Researcher of Optimization of Controlled Processes Department,
ORCID 0000-0002-7246-2736, e-mail: aralova@ukr.net
BELOSHITSKIY P.V.[2], DSc (Medicine), Professor
Professor of Biological Faculty
ORCID 0000-0002-6058-3602, e-mail: bilosh827@ukr.net
ZUBIETA-CALLEJA G.[3], M.D. Professor
Director
ORCID: 0000-0002-4283-6514, e-mail: zubieta@altitudeclinic.com
ARALOVA A.A.[1], PhD (Phys&Math)
Researcher of the Department of Methods for Discrete Optimization,
Mathematical Modelling and Analyses of Complex Systems
ORCID 0000-0001-7282-2036, email: aaaralova@gmail.com
[1] Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine,
40, Acad.Glushkov av., 03187, Kyiv, Ukraine
[2] Pavlo Tychyna Uman State Pedagogical University,
2, Sadova str, 20300 Uman . Chercassy distr., Ukraine
[3] High Altitude Pulmonary and Pathology Institute HAPPI-IPPA
Av. Copacabana, Prolongacion # 55, La Paz, Bolivia

AUTOMATED INFORMATION SYSTEM FOR THE EVALUATION OF CLIMBERS' PERFORMANCE UNDER CONDITIONS OF EXTREMELY LOW pO_2 OF INHALED AIR

Introduction. *Currently, as a result of ever-increasing intensity of human activity, unfavorable environment, the need to perform work in various extreme disturbances, significantly increase physical, mental and emotional stress on the human body, leading to pronounced changes in functional systems. Therefore, the task of studying the adaptation of the human body to work in extreme environments is urgent. The work of climbers is a fairly adequate model for studying the combined effects of hypobaric hypoxia and exercise hypoxia. The need to process large amounts of information necessitates the use of modern computer technology that allows the training process in the training of climbers, which would repeatedly, almost in real time to speed up the processing of survey data and accumulate for further use in determining current status and forecasting regulatory reactions of the body to external and internal disturbances.*

The purpose of the paper *is to develop an automated information system of functional diagnostics using the model of regulation of oxygen regimes of the body and its practical application in the study of highly qualified climbers.*

Methods. *Programming methods for creating an automated information system and methods of functional diagnostics.*

Results. *On the basis of the model of regulation of oxygen regimes of the organism the automated information system for functional diagnostics of the persons who are in the conditions of extreme disturbances is constructed. The results of approbation of the offered software for research of group of highly skilled climbers are resulted.*

Conclusions. *The proposed software allows you to use a model of oxygen regimes of the body in real time, i.e. repeatedly accelerates the processing of data obtained during the survey of athletes, allows centralized collection of information for its pre-processing, storage and collective use, allows you to compare the basic parameters characterizing the functional respiratory system during natural sports activities and obtained during ergometric loading.*

Keywords: *methods of functional diagnostics, highly qualified climbers, mathematical model of regulation of oxygen regimes of the organism, human adaptation to work in extreme environment, hypoxibritic hypoxia.*

INTRODUCTION

The urgency of the problem is caused by the social importance and practical relevance, connected with the necessity of human organism's adaptation to ever increasing requirements of extreme loads in the process of labor and sport activity. Mountain climbing is a vivid example of the combined effect of hypoxybaric and hypermetabolic hypoxia in conditions of increased environmental tension and could be considered as a model problem for research of human activity at high altitudes revealing latent organism's reserve capacities. Nowadays the number of publications on this subject is very insignificant and, unfortunately,most written by non-physiologists — one gets the impression that the authors are completely unaware of the studies, carried out since 1929 at the Institute of Physiology named after A.A. Bogomolets of the National Academy of Sciences of Ukraine and at its Elbrus Medical and Biological Station (EMBS) by a member of the Academy of Sciences N.N. Sirotinin, his students and followers, and in particular in the preparation of Soviet and Ukrainian expeditions to Everest, Manaslu, Kanchejunga, Lhotse, Annapurna, Akonkagua and so on.

Conducted by N.N. Sirotinin school long-term study of the problems of hypoxic states in comparative-physiological, evolutionary aspects, in onto- and phylogeny, at all levels of the organism, using adequate modern methods and mathematical modeling was a broad general biological approach to disclose mechanisms of adaptation, impaired functions, development of mountain sickness, reliability of body functioning in extreme conditions, changes in reactivity and resistance. These fundamental studies, which reveal destructive (pathogenic) and constructive (sanogenic) mechanisms of hypoxic conditions in the body, allowed for the first time in the history of the world to discover and substantiate a new highly effective direction — hypoxytherapy, implemented in the mountains, altitude chambers or with various hypoxicators for treatment, prevention, rehabilitation, increasing resistance and performance.

The concepts on the use of step adaptation to hypoxia, on the organism's oxygen regimes and their regulation, on the functional system of breathing were proposed and substantiated; their mathematical modeling was carried out, allowing not only qualitative but also quantitative characterization of various types of hypoxic states and evaluation of their degrees, forecasting changes in the organism's state under conditions of extreme factors, analyzing the role of certain physiological reactions in compensation of oxygen deficiency.

Naturally, there was a need to process large amounts of information obtained in the study of the body's adaptation to hypoxia, increased work capacity, resistance to the extreme factors of space flight, improved sports performance, which necessitated the introduction of information technology.

These works carried out under the leadership of A.Z. Kolchinskaya, P.V. Beloshitsky, Yu.I. Petunin received back in the seventies the support of a member of the Academy of Sciences V.M. Glushkov and this led in turn to the active introduction into practice the ideas of mathematization of medicine and biology, which initiated the transformation of the science of hypoxia from descriptive experimental to accurate and systematic and generalized results.

The conducted studies have been published in a number of monographs, journals and reported at many international congresses, conferences on mountain medicine and physiology [1–17], and Kyiv was considered the "capital of hypoxia" [6].

We should also note from foreign studies the works [18–30]. Separately, a completely unique study [31] related to the determination of blood gases taken from climbers directly on the peak of Mount Everest should be mentioned.

Nowadays the problems of adaptation to mountain conditions, life at high altitude under chronic hypoxia, respiratory physiology, and neurobiology are intensively studied at the High Altitude Pulmonary and Pathology Institute IPPA (La Paz, Bolivia) since 51 years ago. They initiated the World Congresses on High Altitude Medicine and Physiology in 1994. They established the prestigious international award for outstanding researchers of hypoxia problems (Science, Honor, and Truth) and multiple innovative achievements. Their focus with respect to life at high altitude is based on their observations of human physiology and successful life and exercise activities between 3,100m and 4,100m of altitude in the city of La Paz, with over 2.5 million inhabitants. One of their outstanding feats was to carry out a football (soccer) game on top of Mount Sajama at 6,542m in the highest mountain of Bolivia.

Among their most important scientific observations are:

1) The development of the high altitude adaptation formula [32]:

$$Adaptation = \frac{Time}{Altitude},$$

which states that upon arrival to a high altitude location like the city of La Paz (3.600m), it takes around 40 days for the hematocrit to increase to the maximum normal level for optimal life at high altitude. This is a logarithmic increase reaching a high plateau, the most efficient physiological oxygen transport system under chronic hypoxia. It is interesting to point out that going in a reverse way, there is a linear decrease of the hematocrit to the most optimal sea-level value in 20 days.

2) Another point of interest is the Tolerance to Hypoxia Formula [33]:

$$Tl = \frac{Hb}{p_aCO_2},$$

where Tl — tolerance to hypoxia, was developed based on the fact that hemoglobin increases and p_aCO_2 decreases at high altitude. Noteworthy is the fact that tolerance to hypoxia increase with altitude. On the summit of Mt. Everest, humans are six times more tolerant than at sea level. It is possible to conclude that the human organism carries the capability of survival at high altitude even in the highest point of the Earth.

The Acid-Base balance has also been studied where it was defined that a correction factor should be applied for the Van Slyke formula specific for each altitude [34]. They postulated that maintaining the acid-base balance is transcendental for optimal biochemical function at high altitude, where the pH should remain within normal physiologic values.

The Oxygen Transport Triad [35] is formed by 3 factors: 1) The Pneumo-dynamic pump (the lungs) that is a mechanical vacuum pump allowing ventilation for oxygen renewal in the alveoli and carbon dioxide excretion to the environment. 2) The Hemo-dynamic pump (The heart), which is a mechanical liquid pressure pump that moves blood to and from the tissues, and 3) Hemoglobin, the iron-based molecule that transports oxygen and carbon dioxide. The interrelation between these three mechanisms allows for a most energy-efficient system of survival at high altitude during acute and chronic hypoxic exposure. The Pneumo-dynamic pump plays the most important role of adaptation to acute hypoxia, along with the Hemo-dynamic pump. With adaptation to chronic hypoxia, hemoglobin releases the extra load upon the two pumps, as the energy consumption is high with their increased work.

The hypothesis that man can adapt to live in the hypoxic levels of Mt. Everest [36], continues to be proved with successful climbs. Initially, it was thought that man could not reach the summit without supplemental oxygen. Messner and Habeler were the first to climb Mount Everest breathing only ambient air without an oxygen mask. Messner then climbed all 14, + 8000m mountains without oxygen.

However, these studies, for obvious reasons, are connected with the study of blood: respiratory gas tension, acid-base balance, blood lactate. The works [37, 38] consider the limiting role of the respiratory system. At the same time, the information about the maximum oxygen consumption [39–44], as an index that characterizes the cardiorespiratory system capacity criterion of aerobic capacity, is still relevant when making decisions about the possibility of carrying out extreme loads by the organism. It is believed that the Maximum Oxygen Consumption (MOC) is the factor influencing and limiting the ability to perform in different sports; for climbers, such studies, in particular, were carried out by employees of the Bogomolets Institute of Physiology of the National Academy of Sciences of Ukraine, the National University of Physical Education and Sports of Ukraine and the Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine [45, 46].

In this regard, the methodology of screening the climbers, which was used at the Elbrus medical-biological station when forming teams for climbing the Himalayan eight-thousanders, is of special interest [47–49].

A.Z. Kolchinskaya's concept of regulation of oxygen regimes of human organisms was the theoretical basis for research and decision-making on the formation of expedition composition [50]. According to this concept, two groups of parameters are interconnected in the organism: oxygen transport rates and its partial pressures and stresses at the main stages of its path (lungs, alveoli, arterial, and mixed venous blood). The analysis of combinations of these two groups of parameters allows us to objectively characterize the function of the body's supply system quantitatively and qualitatively.

This approach allows obtaining a general characteristic of gas homeostasis using a minimum of indicators: 1) to give its detailed analysis involving fundamental mechanisms providing respiratory gas transport, 2) to make a diagnosis of the main syndromes related to the disorder of gas transport function, 3) to give an oxygen "portrait" of the organism and its dynamics under various functional states, 4) to assess the organism's ability to recover after external and internal disturbing influences. Systematic accumulation of data, their systematization with subsequent processing and analysis provide discernibility and objectivity of characteristics of a large number of the examined athletes, making it possible to trace the dynamics of changes in the indicators during the annual training cycle by periods of training (transition, basic, competitive), long-term training, increase of sportsmanship, in the age aspect, allowing the establishment of relationships between individual indicators, conduct their differentiation by kinds of sports, trace the differentiation by kinds of sports, trace the dynamics of changes in the main indices in the period after the cessation of active sports activity.

In a healthy individual the oxygen parameters, indicators of efficiency and economy of oxygen regimes of the body, as well as parameters characterizing the production, accumulation, and transport of carbon dioxide, indicators of the internal environment of the body, its acid-base state, and others prove to be so representative that they can be used as normative for a given age, sex, training level, and type of sport. Deviations of oxygen parameters and indicators of economy and efficiency of oxygen regimes from these standards can be used to determine: 1) an objective characteristic of changes in the functional state of the organism, 2) trace the dynamics of this state in the process of athletes' preparation for competitive activity, 3) competitive activity itself and recovery period. It is likewise useful in the process of recovery and rehabilitation of athletes after injuries.

Estimation of general fitness level and adaptation degree of athletes to heavy loads and to oxygen deficiency requires detailed characterization of a complicated process of oxygen delivery to working muscles, the degree of compliance of oxygen delivery, and carbon dioxide excretion process with the metabolic demand of tissues. Such characteristics cannot be given without labor-consuming calculations of some oxygen parameters and functional indices. That was possible due to mathematical models of the respiratory system and computer software.

When analyzing the oxygen regimes of the organism and the criteria of its functional state, a synchronous determination of more than twenty separate indices characterizing the state of the respiratory system is assumed. These indicators include parameters of external respiration, oxygen-transport function of blood, hemodynamic system, and gas exchange. The calculation indices, obtained on their basis allow the characterization of the activity of functional systems of the organism and assess the function of oxygen supply of the organism from the point of view of economy, efficiency, tension, speed, intensity of oxygen delivery in separate sections of its transport within the organism, in a total amount of about one hundred indices.

The mathematical model of mass transfer of gases in human and animal organisms, which was based on the concept of regulation of oxygen regimes of the organism [50], was a reliable tool for characterizing the functional state of athletes in various activities. The role of load hypoxia in resolving the conflict situation between the cardiac and skeletal muscles in the fight for oxygen was investigated using a mathematical model of the breathing system with optimal control.

Calculations of oxygen parameters, carbon dioxide parameters, functional indices additionally allowed the characterization of the activity of functional systems of the organism, the evaluation of the state of oxygen supply system by the rate and intensity of its delivery to lungs and alveoli, the oxygen transport through arterial and venous blood, the tissue oxygen consumption, the efficiency of staged oxygen delivery, and the tension and economy of oxygen modes of the organism.

A sufficiently complete review and analysis of existing developments on this topic is given in [51]. Further development was presented in [52] the automated information system of functional diagnostics of athletes, which included software capabilities existing at that time and its modification for mountaineers [46]. Further development of works in this direction was the study of the dynamics of the parameters of self-organization of the respiratory system of mountain rescuers during short- and medium-term adaptation [53] in the conditions of the middle altitude

The purpose of the work is to create a modern automated information system for functional diagnostics of athletes using the organism oxygen regimes (OOR) model, which would allow:

1) to significantly accelerate the processing of data obtained during the examination of athletes;

2) to centralize the accumulation of information for its pre-processing, storage, and collective use;

3) to create an algorithmic apparatus to provide evidence of scientific provisions, development of options for optimization of decisions on the assessment of athletes' prospects;

4) to implement the diagnostic algorithm of functional state assessment of athletes developed earlier.

SOFTWARE PACKAGE STRUCTURE

For clarity of presentation of the data and their operative processing by means of the correlation analysis, the most informative twenty indicators characterizing the functional condition of the respiratory system were chosen. A computerized model of the respiratory system is constructed. The proposed software allows you to build these model characteristics on the basis of the calculated data. Accordingly, the data stored them in memory with output to external media to create model characteristics for athletes of different ages and training levels. It permits for specialization in certain disciplines of cyclic sports and strength martial arts.

The software works with the implementation of two workstations: a laboratory technician and a medical professional. Such a division was due to the fact that collecting data during the examination requires a set of specific knowledge and skills. Initially such data acquisition knowledge and skills are not necessary, however, the laboratory blood tests performed not directly at the workplace. Indicators characterizing the state of the external and alveolar respiration systems, cardiac activity, circulatory system are registered by devices directly in the process of examination and can be used immediately as initial data.

The system works as follows. After logging into the system at the Laboratory technician's workstation, general data about the subject is entered: such as last name, age, sex, sport, qualification, height, and weight. The system automatically forms groups according to age or qualification. Then, the requests for the examination of a certain group on a certain date are generated on the Automated Workplace (AWP) of the Medic (Fig. 1). Based on the survey requests, general environmental data is collected — barometric pressure, partial pressure of water vapor, altitude, ambient temperature etc. Blood samples are also taken for analysis. Then the respiratory, circulatory, and cardiac systems are examined at rest and under various loads (bicycle ergometry or step tests are possible), depending on the objective set for the researcher. After the examination of one person is completed, the information obtained by the laboratory and the information obtained after performing special performance tests is also entered into the database.

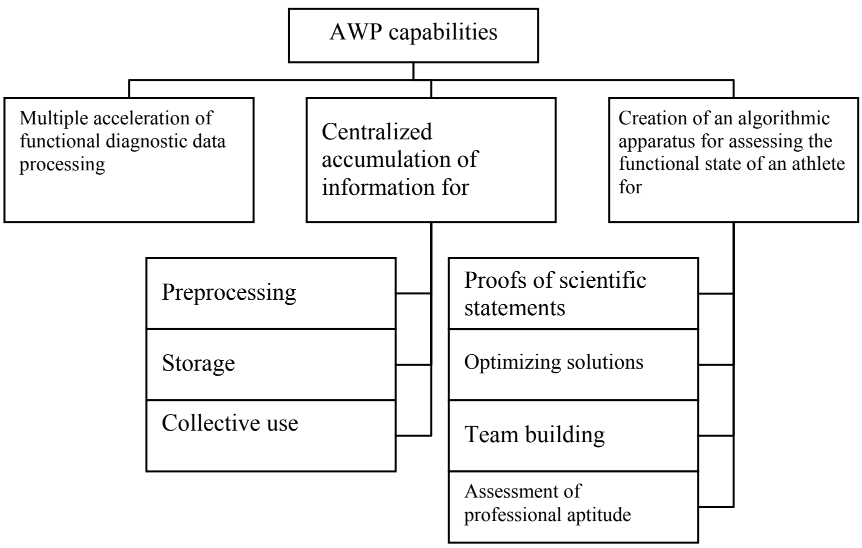


Fig. 1. AWP capabilities

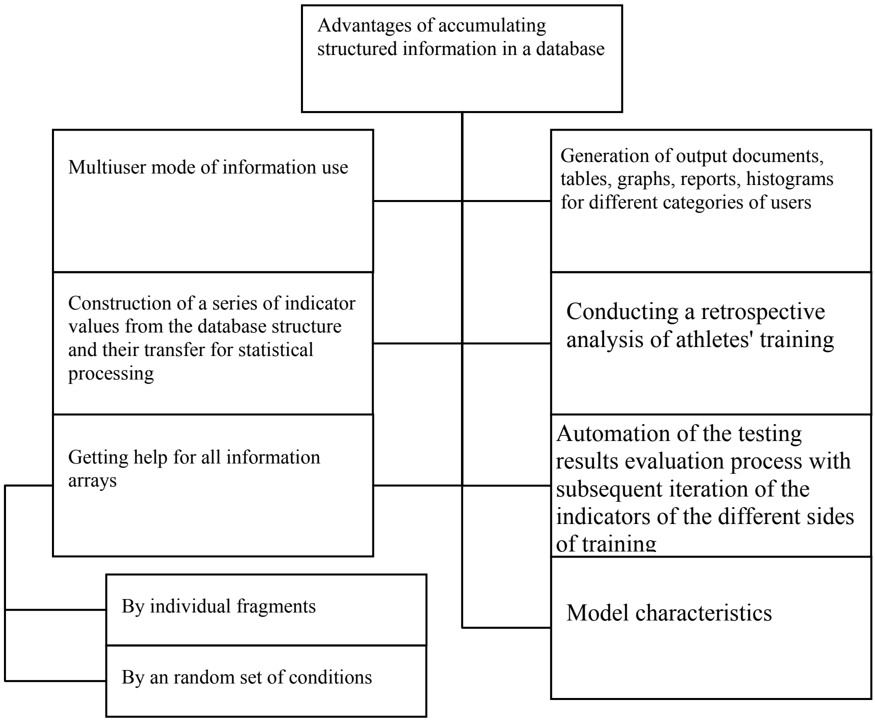


Fig. 2. Benefits of accumulating structured information in a database

Special performance tests were developed for different sports — speed and power, cyclic, technically difficult sports, martial arts, sports games, mountain climbers, and alpinists. This allows for a comparison of the main parameters characterizing the functional breathing system during natural sports activity and those obtained during ergometric bicycle exercise. When all the initial data necessary for the calculation are entered, the calculation of the indicators of oxygen regimes of the organism and their distribution into groups, which correspond to different parts of the respiratory system, takes place. These are functional indices of speed, intensity, the efficiency of staged oxygen delivery; indices characterizing the economy of the respiratory system and blood circulation, as well as parameters characterizing the hypoxic state of the organism. These operations can be repeated several times within one application. In this case, if the weight or height during a series of examinations has changed, the system allows you to enter the changed values in the database while maintaining the previous values. On Fig. 2 are shown benefits of accumulating structured information in a database.

The general scheme of the software package and the OOR algorithm is shown in Fig. 3.

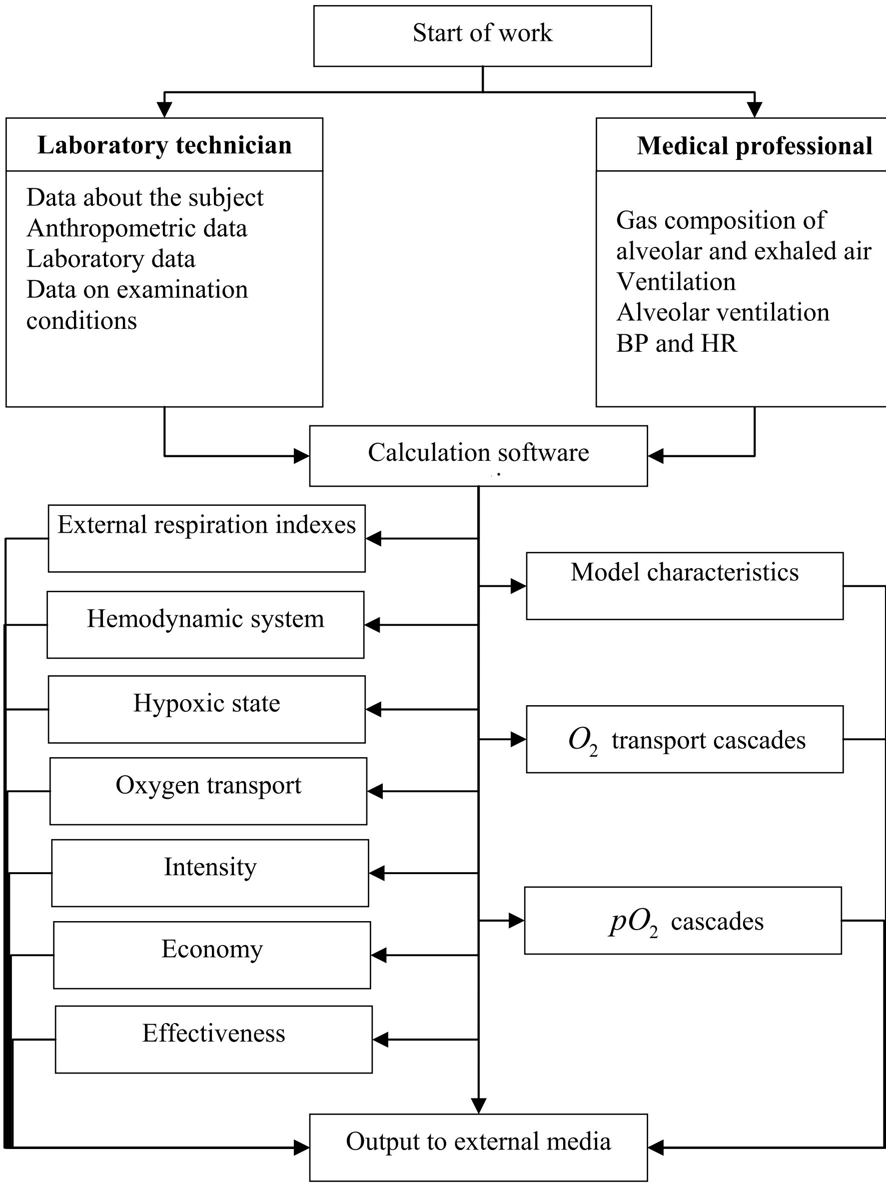


Fig. 3. General scheme of OOR operation

ALGORITHMS OF INDIVIDUAL AWP BLOCKS

The software is divided into separate blocks, each of which solves its own task. Let's describe each of them in detail.

The central block is the Main menu, which is defined by the tasks of the user's workstation. All the blocks of Help, Info and Signing Out are common for both workstations. In other blocks there is a division depending on the tasks, which are solved at this automated workstation. The functions of each block are grouped in Table 1.

Table 1. **Individual block functions**

Block name	Functions
Entering regular data	provides correct input of regular data (full name, sports type, qualification, height, weight). The results are recorded in the database (DB).
Entering personal data	provides correct input of personal data (heart rate, breathing rate etc.). The results are recorded in the database
Survey data entry	provides correct input of the examination data (barometric pressure, gas content in the air etc.). Results are recorded in the database
Entering initial data	provides the correct data input about the examined group, about the system of examination, about the organism state in different moments of examination, including checks on data compatibility. It is carried out both on the level of the software part of the system and on the level of the database management system. The results are recorded in the database and subsequently used to calculate the characteristics
Saving initial data	provides saving the data to the database. Saving takes place automatically and at the user's command. The input is a request to save the original data.
Solving a mathematical task	builds and performs calculations.
Result formation	provides graphical and tabular representation of the results of the work.
Saving a report:	if the user wishes, a report on the performed work can be made. The input is a request to save data about the person being examined, where to save it.
Help for the user	from any task of the system gives contextual help, instructions and advice to the users. The input is a request for help to the user.
Creator Info	from any system task outputs the creator information with contact information. Input is a request for information about the creators
Signing Out	Provides correct sign out, saving parameters, current data via a dialog with the user. The input is a request to sign out of the system.

Let's dwell in more detail on the possibilities of correction and additional data verification. Experience shows that the largest number of failures in calculations and system malfunctions occurs not because of incorrectly working program, but because of an error during data entry. In accordance with the above, the developers have created a robust mechanism for checking the correctness of data, both numeric and lowercase. To do this, various methods of interaction with the user, such as input error messages, convenient structuring, templates for entering string and numeric data, drop-down lists, switches etc are used.

Then, in the dialog box we enter data about the general environmental conditions in which the survey takes place. The given variant of the dialog box assumes entering the known reduction factors to the conditions $BTPS$ and $STPD$, but the variant of calculating these indicators is also provided.

The individual initial data from the examination (minute respiratory volume, alveolar ventilation, exhaled and alveolar gas composition, blood pressure, and heart rate) are then entered.

Further, after entering all the initial data for the group, the calculation of indices of speed, intensity, efficiency of stepwise oxygen delivery, indices characterizing economy of respiratory and circulatory system, parameters characterizing hypoxic state of the organism are made, transport by arterial and mixed venous blood and economy of the system of external respiration and hemodynamics.

RESULTS AND DISCUSSION

In order to study the adaptation of climbers to the reduced partial pressure of oxygen in the inhaled air, the group of 46 climbers was examined. According to the results, it was divided into two groups: sport masters (20 persons) and candidates to sport masters (26 persons). The examination was carried out in natural conditions at an altitude of 2100 m above sea level and in an altitude chamber at an "altitude" of 7500 m above sea level. Cycloergometric load with gradually increasing power was carried out. The minute respiratory volume, gas composition of exhaled and alveolar air, respiratory rate, heart rate, blood pressure, anthropometric parameters, hemoglobin and acidity of arterialized blood were measured.

Using the above-described AWP the indices characterizing the systems of external respiration, blood circulation, efficiency, intensity, economy of oxygen regimes of the body, the parameters of hypoxic state were calculated.

Physical performance of climbers was assessed by the work performed and power developed, the rate of oxygen consumption and carbon dioxide output, the rate of staged oxygen delivery, partial pressure and tension of respiratory gases in the alveolar space, arterial and mixed venous blood. The degree of tissue hypoxia was assessed by the presence of oxygen debt, shifts in the acid-base state of blood, changes in blood acidity, presence of lactate.

The test of special work capacity of climbers was also carried out. The obtained indices were also compared with similar indices obtained when examining these groups on the plain (lowlands) [49]. The results of the examinati on are presented in the Table 2.

Table 2. **The results of the examination**

Indicator	**Load**			
	1,7 Wt/kg		**2,7 Wt/kg**	
	2100 m	**7500 m**	**2100 m**	**7500 m**
Respiratory volume per minute, l/min	57,0±1,8	110,3±3,4	72,2±2,3	137,9±5,0
Respiratory rate, breath/min	22,7±2,1	34,25±3,2	28,6±1,9	49,96±4,1
Respiratory volume, l	2,51±0,1	3,22±0,12	2,52±0,23	2,74±0,14
Ventilatory equivalent,	28,57±1,6	76,39±2,9	29,76±1,9	83,37±1,8
Respiratory cycle oxygen effect, ml/r.c.	87,9±2,1	42,15±3,3	84,82±1,8	33,1±2,7
Partial pressure of oxygen in the alveolar space, mm Hg	76,13±2,0	36,94±1,6	75,47±1,6	38,83±2,1
Partial pressure of carbon dioxide in alveolar space, mm Hg	31,9±0,8	13,34±0,6	33,88±1,0	13,24±0,4
Oxygen consumption rate, l/min	1,995±0,093	1,444±0,106	2,426±0,104	1,654±0,141
Heart Rate, r/m	105,2±3,1	139,7±2,8	123,8±4,2	149,9±3,2
Systolic output, ml	133±9	119,13±2,0	134,9±5	118,8±1,8
Cardiac output, l/min	13,965±0,35	16,71±0,205	16,739±0,48	17,72±0,345
Hemodynamic equivalent	7,0±0,2	11,57±0,7	6,75±0,2	10,7±0,9
Cardial cycle oxygen effect, ml/c.c.	18,96±1,1	10,33±0,6	19,59±0,9	11,03±0,85
Hemoglobin content, g/l	154,2±3	159,2±2	154,2±3	159,2±2
Arterial blood oxygen content, ml/l	188,8±2,4	138,6±1,8	184,2±3,1	129,9±1,1
Arterial blood oxygen saturation, %	90,1±1,3	64,0±0,9	89,3±13	59,98±1,1

In order to find out the climbers' adaptability to extremely low partial pressure of oxygen in the inhaled air the performance and functional state of the organism were determined: at the "altitude " of 7500 meters there was carried out a cycloergometric examination in an altitude chamber. The loads were reduced. The first load was 0.85 W/kg, (we do not consider this load because there is nothing to compare it with), the second — 1.7 W/kg, the third — 2.7 W/kg, i.e. there was work that was not more intensive than moderate at sea level and at 2000 meters (oxygen consumption at these loads and at 2100 meters was less than 50% of MOC). More than half of the examined climbers carried a load of 2.7 W/kg. Those climbers who were able to withstand it performed it either at the AMT level (anaerobic metabolic threshold) or slightly to the right beyond it, i.e. third-degree load hypoxia was manifested.

The data obtained show that breathing and blood circulation at the "altitude" of 7500 m during low-intensity load at sea level becomes less economical. At the simulated altitude of 7500 m the second load 1.7 W/kg, (the first load at 2100m), was performed with an oxygen consumption 551 ml/min lower than at 2100m (differences are not significant, $p>0.05$). The climbers consumed on average 1444. ± 107 ml/min of oxygen. At the same time, each liter of consumed oxygen provided 76.4 ± 94.1 l/min of ventilated air, which is 47.81 l/min more than at 2100 m altitude. The consumption of 1 L of oxygen required 11.57±0.8 L of circulating blood, which was 4.57 L more than at 2100 m for the same workload. Systolic volume was lower than at 2100 m and was 119.13+3.4 ml, and cardiac output was 2.745 L/min higher than during the same work at 2100 m. At a load of 1.7 W/kg, the arterio-venous oxygen difference was quite high (132.5 ml/l ± 3.0 ml/l), the oxygen content in mixed venous blood decreased to 53.6 ml/l ± 2.3 ml / l, its saturation with oxygen up to 25% ± 1.6%, oxygen tension up to 13 mm Hg. At a load of 2.7 W/kg, the arteriovenous oxygen difference increased to 145.1 ml/l ± 4.2 ml/l, the oxygen content was 28.1 ml/l ± 0.9 ml/l, the saturation of mixed venous blood with oxygen was only 13.3% ± 0.6%, oxygen tension in mixed venous blood was 12.0 ± 0.4 mm Hg. At the "altitude" of 7500 m, the coefficient of oxygen utilization from the blood decreased, as evidenced by the decrease in the arterio-venous difference in oxygen to 86.4 ml/l ± 2.1 ml/l at a load of 1.7 W/kg and to 72.9 ml /l ± 1.0 ml/l. Note that at a load of 1.7 W/kg, the oxygen content in the mixed venous blood was 52.4 ml/l ± 2.4 ml/l, saturation 24.2% ± 0.6%. At a load of 2.7 W/kg, the oxygen content in the mixed venous blood was 36.7 ml/l ± 3.1 ml/l and 17.2% ± 0.9%, respectively. The oxygen tension in the mixed venous blood at loads of 1.7 W/kg and 2.7 W/kg was 12 ± 0.5 mm Hg, respectively and 13 ± 0.9 mm Hg.

In [49] the impossibility of maintaining a high rate of oxygen consumption in conditions of extremely low ambient air is explained, on the one hand, by the decreasing rate of oxygen diffusion from the blood capillaries into cells and mitochondria due to a significantly reduced gradient between the blood and the intracellular environment, on the other hand, oxygen tension in arterial and mixed venous blood below a critical level causes a significant decrease in tissues, which directly reduces the rate of oxidative processes.

During a load of 2.7 W/kg (the third load at an "altitude" of 7500 m and the second load at 2100 m), oxygen consumption decreased and the efficiency of external respiration decreased — each liter of oxygen at this load provided 84.02 ±4.1 l/min of ventilated air, 40 l/min more than at the same load at 2100 m. The respiratory volume did not increase significantly, but the respiratory rate increased, which was 43.79 ± 6 breath/min. Respiratory coefficient began to exceed unity, arterio-venous difference decreased markedly, and hemodynamic equivalent increased. Oxygen tension in arterial blood was the same as during lower intensity exercise performed at the same altitude, and oxygen tension in mixed venous blood was 2-4 mm higher than its values during previous exercise. Low arterio-venous difference and increased pO_2 in the mixed venous blood indicated the decrease of oxygen utilization from the blood, the reason of which was the decrease of its tension in the tissues to below the critical level. Metabolic acidosis became more expressed: at the "altitude" of 7500 m pH was 0.11 lower than at rest; while pO_2 in alveolar air was only 16.58±1.6 mm Hg, i.e. 30.57 mm Hg lower than at sea level.

Estimation of general physical fitness level and adaptation degree of athletes to heavy loads and to oxygen deficiency requires detailed characterization of the complex process of oxygen delivery to the working muscles, the degree of compliance of oxygen delivery and carbon dioxide excretion process with the metabolic demand of tissues. Such characterization cannot be given without labor-consuming calculations of some oxygen parameters and functional indices, which turned out to be possible due to created mathematical models of respiratory system and software for computer.

The mathematical model of mass transfer of gases in human and animal organisms, which was based on the concept of regulation of oxygen regimes of the organism [50], was a reliable tool for characterizing the functional state of athletes in various activities. The role of load hypoxia in resolving the conflict between the cardiac and skeletal muscles in the struggle for oxygen was investigated using a mathematical model of the respiratory system with optimal control.

Calculations of oxygen parameters, carbon dioxide parameters, functional indices additionally allowed to characterize the activity of functional systems of the organism, to estimate the state of oxygen supply system by the rate and intensity of its entrance into the lungs and alveoli, oxygen transport by arterial and venous blood, tissue consumption, efficiency of staged oxygen delivery, tension and economy of oxygen regimes of the organism.

Reduced oxygen consumption rate, decreased arterio-venous oxygen difference, increased oxygen debt and blood lactate content, decreased *pH*, increased respiratory coefficient, low O_2 stress in arterial and mixed venous blood indicated that that load hypoxia during work with 2.7 W/kg at the "altitude" of 7500 m became uncompensated, which made it impossible to perform heavy loads at this altitude, i.e. at 7500 m the load of 2.7 W/kg can be considered the limit for climbers. Out of 60 people who were examined at this altitude, nine climbers were not able to carry these loads for more than one minute. Large shifts of *pH* after loading at the "altitude" of 7500 m in the pressure chamber indicated insufficient function of compensatory mechanisms and insufficient adaptation of these climbers to strenuous muscular activity at high altitudes.

Let us also note the following. If we analyze separately the role of the external respiratory system and the circulatory system in adaptation to the combined effects of hypoxybaric hypoxia and load hypoxia (Figure 4, 5), it appears that adaptation is mainly due to the external respiratory system, both when the load increases and when the partial pressure of oxygen in the inhaled air decreases.

At the altitude of 2100 m when the load increased from 1.7 W/kg to 2.7 W/kg, respiratory volume increased by 26 %; at the "altitude" of 7500 m when the load increased from 1.7 W/kg to 2.7 W/kg, ventilation increased by 25 %. At the same time the main role in increasing ventilation was performed by 26 % and 45 % increase in respiratory rate, respiratory volume at the altitude of 2100 m increased by 28 %, at the "altitude" of 7500 m respiratory volume decreased by 17 %. If we compare the dynamics of the external respiratory system indices during the same load of 1.7 W/kg, it appears that the main role in the increase of MRV by 142 % was played by the increase of respiratory frequency by 120 %, while the respiratory volume increased by only 9 % (Figure 6–8).

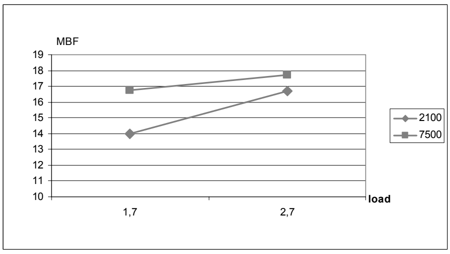


Fig. 4. The compensatory role of the external respiratory system in adaptation to the combined effects of hypoxybaric hypoxia and load hypoxia

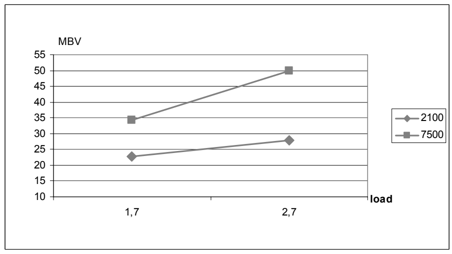


Fig. 5. The compensatory role of the circulatory system during adaptation to the combined effects of hypoxybaric hypoxia and load hypoxia

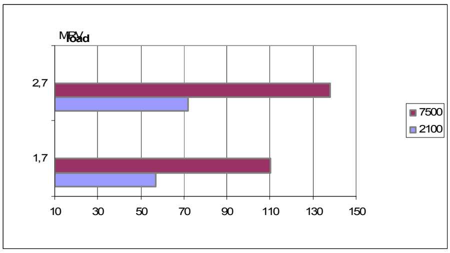


Fig. 6. Minute Respiratory Volume

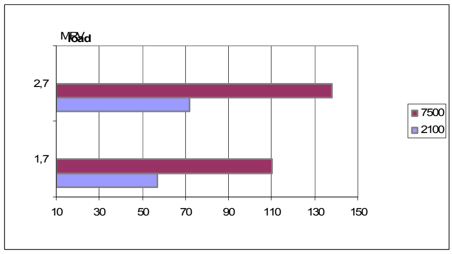


Fig. 7. Respiratory Erequency

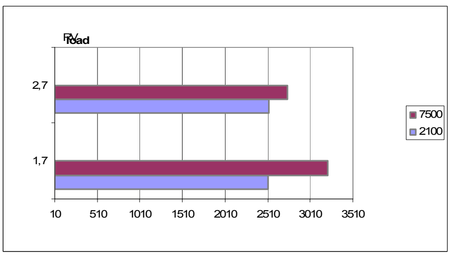


Fig. 8. Respiratory Volume

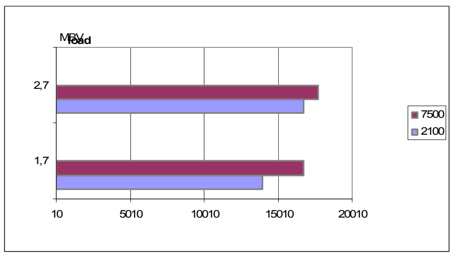


Fig. 9. Systemic Blood Flow

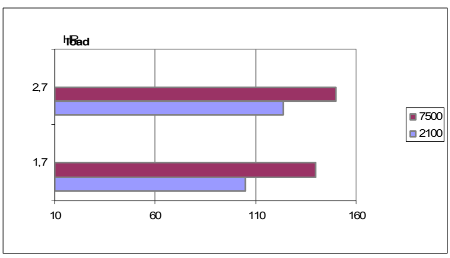


Fig. 10. Heart Rate

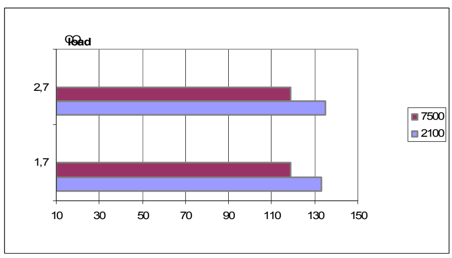


Fig. 11 Respiratory Erequency

If we consider the role of systemic blood flow (SBF) and its components — stroke volume and heart rate (Figure 9–11), we find that SBF increased insignificantly at the altitude of 2100 m when the load was increased from 1.7 W/kg to 2.7 W/kg, and at the first load at the "altitude" of 7500 m — only by 19 %, at the second load at 7500 m the increase was even less — only by 6 %. When performing the same load of 1.7 W/kg at 2100 and 7500 m, there was a 26% increase in blood flow. At the same time, the main role in the increase of the systemic blood flow is played by the increase of heart rate: by 17 % at the "altitude" of 7500 m compared to 2100 m, with the increase of load at the altitude of 2100 m the heart rate increased by 33 %, with the increase of load at the "altitude" of 7500 m there was an increase of heart rate by 21 %. As for the systolic output, the study showed that with increasing load, the value of this parameter decreased both at 2100 m altitude by 12 % and at 7500 m altitude by 13 %, with the values of these figures being practically equal.

CONCLUSION

The proposed software allows using the model of organism's oxygen regimes practically in real time, because it accelerates processing of the data obtained in the process of examination of sportsmen and allows centralized accumulation of information for its pre-processing, storage and collective use, allows to compare the main parameters characterizing the functional system of breathing during natural sports activity and those obtained during cycloergometric exercise, recovery, in the annual training cycle, during the Olympic cycle, during the whole sports activity and after it, during the recovery and rehabilitation after severe injuries when training activity was suspended, to study the processes of adaptation of the athlete's organism to training loads.

REFERENCES

1. Sirotinin N.N. Living on the heights and ailments on the heels. Kiev: Edition of the Academy of Sciences of the URS,1939, 222 p. (In Ukrainian)
2. Biloshytsky P.V. Hypoxia. Encyclopedia of Successful Ukraine Kyiv: Institute of Encyclopedic Research NAS of Ukraine. 2006, **5**, pp. 625–626. (In Ukrainian)
3. Beloshitskiy P.V. Study of the problems of sports medicine at the Elbrus medical and biological station. Sports medicine. 2008, 1, pp. 83–94. (In Russian)
4. Kolchinskaya A.Z. On the classification of hypoxic conditions. Patol. physiology and experimental therapy. 1981, no. 4, pp. 3–10. (In Russian)
5. Beloshitskiy P.V. Chronicle of biomedical research in the Elbrus region (1929–2006). Kyiv. 2014. Ukrainian Academy of Sciences. 550 p. (In Russian)
6. Seredenko M.M., Nazarenko A.I, Rozova. K.V. The capital of hypoxia. Kyiv, 2002. Print line. 24 p (In Ukrainian)
7. First world congress of high altitude medicine and physiology.La Paz-Bolivia. 1994, p.
8. Second world congress of high altitude medicine and physiology. *Acta Andina.* 1996, Cusco-Peru.
9. Materials to the II Congress of Pathophysiologists of Ukraine, dedicated to the 100th anniversary of the day of the Birth Day of Ac. M.M. Syrotinin. *Physiol. journal.* v.42, no 3–4. (In Ukrainian)
10. The 3rd World Congress on Mountain Medicine and High Altitude Physiology and the 18th Japanese Symposium on Mountain Medicine, May 20th–24th. 1998. Matsumoto, Japan.
11. IV World Congress on Mountain Medicine and High Altitude Physiology. *High Altitude medicine and biology*. 2000, 1, N3, Arica, Chile.
12. V World Congress on Mountain Medicine and High Altitude Physiology. *High Altitude medicine and biology.* 2002, 3, N1, Barcelona, Spain.

13. Milledge J.S. VI World Congress on Mountain Medicine & High Altitude Physiology, Xining, Qinghai, and Lhasa, Tibet, August 12–18, 2004. High Alt Med Biol. 2004 Winter;5(4):457-64. doi: 10.1089/ham.2004.5.457.
14. VI World Congress on Mountain Medicine and High Altitude Physiology, Xining, Qinghai, and Lhasa, Tibet, August 12–18, 2004. High Altitude medicine and biology. 2004. 5, N2
15. Anholm J.D. The VIIIth World Congress of High Altitude Medicine and Physiology, Arequipa, Peru, August 8–12, 2010. High Alt Med Biol. 2010 Winter;11(4):381-4. doi: 10.1089/ham.2010.1061.
16. West J.B. World Congresses of Mountain Medicine and High Altitude Physiology, 1994–2002 *High Altitude Medicine & Biology.* V. 3, No. 1. Published Online: 6 Jul 2004 https://doi.org/10.1089/152702902753639478
17. 7th Chronic Hypoxia Symposium, Feb 23 – Mar 2. 2019 La Paz–Bolivia. Dedicated to the Late Danish Prof. PoulErikPaulev p. 24. URL:https://zuniv.net/ symposium7/Abstracts7CHS.pdf
18. Catron T.F., Powell F.L., West J.B. A strategy for determining arterial blood gases on the summit of Mt. Everest. BMC Physiol. 2006 Mar 8;6:3. doi: 10.1186/1472-6793-6-3.
19. West J.B., Hackett P.H., Maret K.H. Milledge J.S., Peters R.M. Jr., Pizzo C.J., Winslow R.M. Pulmonary gas exchange on the summit of Mount Everest. *J Appl Physiol Respir Environ Exerc Physiol.* 1983 Sep;55(3):678-87. doi: 10.1152/jappl.1983.55.3.678.
20. Malconian M.K., Rock P.B., Reeves J.T., Cymerman A., Houston C.S. Operation Everest II: gas tensions in expired air and arterial blood at extreme altitude. *Aviat Space Environ Med.* 1993 Jan;64(1):37–42.
21. West J.B. Lactate during exercise at extreme altitude. *Fed Proc.* 1986 Dec;45(13):2953-7.
22. West J.B. Tolerance to severe hypoxia: lessons from Mt. Everest. *Acta Anaesthesiol Scand Suppl.* 1990;94:18-23. doi: 10.1111/j.1399-6576.1990.tb03216.x.
23. West J.B. Acclimatization and tolerance to extreme altitude. *J Wilderness Med*. 1993 Feb;4(1):17–26. doi: 10.1580/0953-9859-4.1.17.
24. West J.B. Limiting factors for exercise at extreme altitudes. *Clin Physiol.* 1990 May;10(3):265-72. doi: 10.1111/j.1475-097x.1990.tb00095.x.
25. West J.B. Climbing Mt. Everest without oxygen: an analysis of maximal exercise during extreme hypoxia. *Respir Physiol.* 1983 Jun;52(3):265-79. doi: 10.1016/0034-5687(83)90085-3.
26. Wagner P.D., Wagner H.E., Groves B.M., Cymerman A., Houston C.S. Hemoglobin P(50) during a simulated ascent of Mt. Everest, Operation Everest II. *High Alt Med Biol.* 2007 Spring;8(1):32-42. doi: 10.1089/ham.2006.1049.
27. Wagner P.D. Operation Everest II. *High Alt Med Biol.* 2010 Summer;11(2):111-9. doi: 10.1089/ham.2009.1084.
28. Samaja M., Mariani C., Prestini A., Cerretelli P. Acid-base balance and O2 transport at high altitude. *Acta Physiol Scand.* 1997 Mar;159(3):249-56. doi: 10.1046/j.1365-201X.1997.574342000.x.
29. Hoiland R.L., Howe C.A., Coombs G.B., Ainslie P.N. Ventilatory and cerebrovascular regulation and integration at high-altitude. *Clin Auton Res.* 2018 Aug;28(4):423-435. doi: 10.1007/s10286-018-0522-2.
30. Bärtsch P., Saltin B. General introduction to altitude adaptation and mountain sickness. Scand *J Med Sci Sports.* 2008 Aug;18 Suppl 1:1-10. doi: 10.1111/j.1600-0838.2008.00827.x.
31. Grocott M.P., Martin D.S., Levett D.Z., McMorrow R., Windsor J., Montgomery H.E.; Caudwell Xtreme Everest Research Group. Arterial blood gases and oxygen content in climbers on Mount Everest. *N Engl J Med.* 2009 Jan 8;360(2):140-9. doi: 10.1056/NEJMoa0801581.
32. Zubieta-Calleja G.R., Paulev P.E., Zubieta-Calleja L., Zubieta-Castillo G. Altitude adaptation through hematocrit changes. J Physiol Pharmacol. 2007 Nov;58 Suppl 5(Pt 2):811-8. PMID: 18204195.
33. Zubieta-Calleja G.R., Ardaya G., Zubieta N., Paulev P.E., Zubieta-Castillo G. Tolerance to Hypoxia [Internet]. Vol. 59, J Fisiol. 2013. p. 65–71. URL: https://zuniv.net/pub/TolerancetoHypoxiaFiziol.pdf

34. Paulev P.E., Zubieta-Calleja G.R. Essentials in the diagnosis of acid-base disorders and their high altitude application. *J Physiol Pharmacol.* 2005 Sep; 56 Suppl 4:155-70. PMID: 16204789.
35. Zubieta-Calleja G., Zubieta-DeUrioste N. The Oxygen Transport Triad in High-Altitude Pulmonary Edema: A Perspective from the High Andes. Int J Environ Res Public Health. 2021 Jul 17;18(14):7619. doi: 10.3390/ijerph18147619. PMID: 34300070; PMCID: PMC8305285.
36. Zubieta-Castillo G., Zubieta-Calleja G.R., Zubieta-Calleja L., Zubieta-Calleja Nancy. Adaptation to life at the altitude of the summit of Everest. *Physiol. Journal*, 1994.
37. Schoene R.B. Limits of human lung function at high altitude. *J Exp Biol.* 2001 Sep;204(Pt 18):3121-7.
38. Schoene R.B. Limits of respiration at high altitude. *Clin Chest Med.* 2005 Sep;26(3):405-14, vi. doi: 10.1016/j.ccm.2005.06.015.
39. Hawkins M.N. Raven P.B., Snell P.G., Stray-Gundersen J., Levine B.D. Maximal oxygen uptake as a parametric measure of cardiorespiratory capacity. *Med Sci Sports Exerc.* 2007 Jan;39(1):103-7. doi: 10.1249/01.mss.0000241641.75101.64.
40. Noakes T.D. Maximal oxygen uptake as a parametric measure of cardiorespiratory capacity: comment. *Med Sci Sports Exerc.* 2008 Mar;40(3):585; author reply 586. doi: 10.1249/MSS.0b013e3181617350.
41. Bassett D.R. Jr., Howley E.T. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc.* 2000 Jan;32(1):70-84. doi: 10.1097/00005768-200001000-00012
42. Feitosa W.G., Barbosa T.M., Correia R. De A., Castro F.A. De S. Maximal oxygen uptake, total metabolic energy expenditure, and energy cost in swimmers with physical disabilities. *International Journal of Performance Analysis in Sport* Volume 19, 2019 Issue 4, pp. 503–516
43. Kang M.-Y., Sapoval B. Prediction of maximal oxygen uptake at high altitude. *European Respiratory Journal* 2016 48: PA1583; DOI: 10.1183/13993003.congress-2016.PA1583
44. Bernardi L., Schneider A., Pomidori. L., Paolucci E., Cogo A. Hypoxic ventilatory response in successful extreme altitude climbers. *European Respiratory Journal* 2006 27: 165-171; DOI: 10.1183/09031936.06.00015805
45. Biloshitsky P.V., Onopchyk Yu.M., Marchenko D.I., Aralova N.I. Mathematic methods for investigation of reliability problem of organism functioning in extreme high mountain conditions *Physiol. Journal*, 2003, pp. 139–143 (In Ukrain)
46. Aralova A.A., Aralova N.I., Beloshitsky P.V., Onopchuk Yu.N. Automated Information System for Functional Diagnostics of Mountaineers. *Sports Medicine* 2008. 1: 163–169.
47. Kolchinskaya. A.Z., Monogarov V.D., Radzievsky P.A., Molchanova N.I. Complex control in mountaineering and its role in the management of the training process. Control of the process of adaptation of the organism of sports-shifts of high qualification: Collection of scientific works. works / Ed. board: D.A. Polishchuk and others Kyiv.KSIPT 1992, pp.122–132 (in Russian)
48. Kolchinskaya A.Z., Monogarov V.D., Radzievsky P.A. Scientific and methodological support for the preparation of Soviet climbers for the traverse of the Kanchenjunga massif. *Theory and practice of physical culture*. 1991.4.10-14. (in Russian)
49. Kolchinskaya A.Z., Beloshitsky P.V., Monogarov V.D., Pivnutel R.V., Radzievsky P.A., Krasyuk A.N., Ivashkevich A.A., Borisov A.N. Physiological performance of climbers in conditions of extremely low pO_2 in inhaled air. *Physiological journal.* 1989, 35, 2, pp. 68–74. (in Russian)
50. Lauer N.V., Kolchinskaya A.Z. About the oxygen organism regime and its regulation. K.: Nauk.Dumka, 1966, pp. 157–200/ (In Russian)
51. Klyuchko O.M., Aralova N.I., Aralova A.A. Electronic automated work places for biological investigations Biotechnologia Acta, K, 2019, V.12, №2, pp. 5–26. https://doi.org/10.15407 /biotech12/02/005
52. Aralova A.A., Aralova N.I., Kovalchuk-Khimyuk L.A., Onopchuk Yu.N., Automated information system for athletes functional diagnostics. *Control systems and machines.* 2008 3: 73–78. (in Russian).

53. Aralova N.I. Mathematical model of the mechanism short- and medium-functional adaptation of breath of persons work in extreme conditions high. *Kibernetika i vyčislitelnaâ tehnika.* 2015, 182, pp. 15–25. (in Russian).

Received 19.01.2022

ЛІТЕРАТУРА

1. Сиротинін М.М.Життя на висотах і хвороба висоти.Київ: Вид-во АН УРСР.1939. 222 с.
2. Білошицький П.В. Гіпоксія. Енциклопедія сучасної України. Інститут енциклопедичних досліджень НАН України **5**. С. 625–626.
3. Белошицкий П.В. Изучение проблем спортивной медицины на Эльбрусской медико-биологической станции.*Спортивна медицина*. 2008. 1. С. 83–94.
4. Колчинская А.З. О классификации гипоксических состояний. *Патол. физиология и эксперим. терапия*.1981.Вып. 4. С. 3–10.
5. Белошицкий П. В. Летопись медико-биологических исследований в Приэльбрусье (1929–2006 гг.).Киев. 2014. Украинская академия наук. 550 с.
6. Середенко М.М., Назаренко А.І., Розова. К.В. Столиця гіпоксії. Київ, 2002. Прінт лайн. 24 с.
7. First world congress of high altitude medicine and physiology.La Paz-Bolivia. 1994
8. Second world congress of high altitude medicine and physiology. *Acta Andina.* 1996. Cusco-Peru.
9. Матеріали II конгресу патофізіологів України, присвяченого 100-річчю від дня народження ак. М.М. Сиротініна. *Фізіологічний журнал* , т.42. № 3–4.
10. The 3rd World Congress on Mountain Medicine and High Altitude Physiology and the 18th Japanese Symposium on Mountain Medicine, May 20th–24th. 1998. Matsumoto, Japan.
11. IV World Congress on Mountain Medicine and High Altitude Physiology. *High Altitude medicine and biology*. 2000. 1. N3. Arica, Chile.
12. V World Congress on Mountain Medicine and High Altitude Physiology. *High Altitude medicine and biology.* 2002. 3. N1. Barcelona, Spain.
13. Milledge J.S. VI World Congress on Mountain Medicine & High Altitude Physiology, Xining, Qinghai, and Lhasa, Tibet, August 12-18, 2004. High Alt Med Biol. 2004 Winter;5(4):457-64. doi: 10.1089/ham.2004.5.457.
14. VI World Congress on Mountain Medicine and High Altitude Physiology, Xining, Qinghai, and Lhasa, Tibet, August 12-18, 2004. High Altitude medicine and biology. 2004. 5. N2.
15. Anholm J.D. The VIIIth World Congress of High Altitude Medicine and Physiology, Arequipa, Peru, August 8-12, 2010. High Alt Med Biol. 2010 Winter;11(4):381-4. doi: 10.1089/ham.2010.1061.
16. West J. B. World Congresses of Mountain Medicine and High Altitude Physiology, 1994–2002 *High Altitude Medicine & Biology.* V. 3, No. 1. Published Online: 6 Jul 2004 https://doi.org/10.1089/152702902753639478
17. 7th Chronic Hypoxia Symposium, Feb 23 – Mar 2 2019 La Paz–Bolivia. Dedicated to the Late Danish Prof. PoulErik Paulev P. 24. URL: https://zuniv.net/ symposium7/Abstracts7CHS.pdf
18. Catron T.F., Powell F.L., West J.B. A strategy for determining arterial blood gases on the summit of Mt. Everest. BMC Physiol. 2006 Mar 8;6:3. doi: 10.1186/1472-6793-6-3.
19. West J.B., Hackett P.H., Maret K.H. Milledge J.S., Peters R.M. Jr., Pizzo C.J., Winslow R.M. Pulmonary gas exchange on the summit of Mount Everest. *J Appl Physiol Respir Environ Exerc Physiol.* 1983 Sep;55(3):678-87. doi: 10.1152/jappl.1983.55.3.678.
20. Malconian M.K., Rock P.B., Reeves J.T., Cymerman A., Houston C.S. Operation Everest II: gas tensions in expired air and arterial blood at extreme altitude. *Aviat Space Environ Med.* 1993 Jan;64(1):37-42.
21. West J.B. Lactate during exercise at extreme altitude. *Fed Proc.* 1986 Dec;45(13):2953-7.
22. West J.B. Tolerance to severe hypoxia: lessons from Mt. Everest. *Acta Anaesthesiol Scand Suppl.* 1990;94:18-23. doi: 10.1111/j.1399-6576.1990.tb03216.x.
23. West J.B. Acclimatization and tolerance to extreme altitude. *J Wilderness Med*. 1993 Feb;4(1):17-26. doi: 10.1580/0953-9859-4.1.17.
24. West J.B. Limiting factors for exercise at extreme altitudes. *Clin Physiol.* 1990 May;10(3):265-72. doi: 10.1111/j.1475-097x.1990.tb00095.x.

25. West J.B. Climbing Mt. Everest without oxygen: an analysis of maximal exercise during extreme hypoxia. *Respir Physiol.* 1983 Jun;52(3):265-79. doi: 10.1016/0034-5687(83)90085-3.
26. Wagner P.D., Wagner H.E., Groves B.M., Cymerman A., Houston C.S. Hemoglobin P(50) during a simulated ascent of Mt. Everest, Operation Everest II. *High Alt Med Biol.* 2007 Spring;8(1):32-42. doi: 10.1089/ham.2006.1049.
27. Wagner P.D. Operation Everest II. *High Alt Med Biol.* 2010 Summer;11(2):111-9. doi: 10.1089/ham.2009.1084.
28. Samaja M., Mariani C., Prestini A., Cerretelli P. Acid-base balance and O2 transport at high altitude. *Acta Physiol Scand.* 1997 Mar;159(3):249-56. doi: 10.1046/j.1365-201X.1997.574342000.x.
29. Hoiland R.L., Howe C.A., Coombs G.B., Ainslie P.N. Ventilatory and cerebrovascular regulation and integration at high-altitude. *Clin Auton Res.* 2018 Aug;28(4):423-435. doi: 10.1007/s10286-018-0522-2.
30. Bärtsch P., Saltin B. General introduction to altitude adaptation and mountain sickness. Scand *J Med Sci Sports.* 2008 Aug;18 Suppl 1:1-10. doi: 10.1111/j.1600-0838.2008.00827.x.
31. Grocott M.P., Martin D.S., Levett D.Z., McMorrow R., Windsor J., Montgomery H.E.; Caudwell Xtreme Everest Research Group. Arterial blood gases and oxygen content in climbers on Mount Everest. *N Engl J Med.* 2009 Jan 8;360(2):140-9. doi: 10.1056/NEJMoa0801581.
32. Zubieta-Calleja G.R., Paulev P.E., Zubieta-Calleja L., Zubieta-Castillo G. Altitude adaptation through hematocrit changes. J Physiol Pharmacol. 2007 Nov;58 Suppl 5 (Pt 2):811-8. PMID: 18204195.
33. Zubieta-Calleja G.R., Ardaya G., Zubieta N., Paulev P.E., Zubieta-Castillo G. Tolerance to Hypoxia [Internet]. Vol. 59, J Fisiol. 2013. p. 65–71. URL: https://zuniv.net/pub/TolerancetoHypoxiaFiziol.pdf
34. Paulev P.E., Zubieta-Calleja G.R. Essentials in the diagnosis of acid-base disorders and their high altitude application. *J Physiol Pharmacol.* 2005 Sep;56 Suppl 4:155-70. PMID: 16204789.
35. Zubieta-Calleja G., Zubieta-DeUrioste N. The Oxygen Transport Triad in High-Altitude Pulmonary Edema: A Perspective from the High Andes. Int J Environ Res Public Health. 2021 Jul 17;18(14):7619. doi: 10.3390/ijerph18147619. PMID: 34300070; PMCID: PMC8305285.
36. Zubieta-Castillo G., Zubieta-Calleja G.R., Zubieta-Calleja L., Zubieta-Calleja Nancy. Adaptation to life at the altitude of the summit of Everest. *Physiol. Journal*, 1994.
37. Schoene R.B. Limits of human lung function at high altitude. *J Exp Biol.* 2001 Sep;204(Pt 18):3121-7.
38. Schoene R.B. Limits of respiration at high altitude. *Clin Chest Med.* 2005 Sep;26(3):405-14, vi. doi: 10.1016/j.ccm.2005.06.015.
39. Hawkins M.N. Raven P.B., Snell P.G., Stray-Gundersen J., Levine B.D. Maximal oxygen uptake as a parametric measure of cardiorespiratory capacity. *Med Sci Sports Exerc.* 2007 Jan;39(1):103-7. doi: 10.1249/01.mss.0000241641.75101.64.
40. Noakes T.D. Maximal oxygen uptake as a parametric measure of cardiorespiratory capacity: comment. *Med Sci Sports Exerc.* 2008 Mar;40(3):585; author reply 586. doi: 10.1249/MSS.0b013e3181617350.
41. Bassett D.R. Jr., Howley E.T. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc.* 2000 Jan;32(1):70-84. doi: 10.1097/00005768-200001000-00012
42. Feitosa W.G., Barbosa T.M., Correia R. De A., Castro F.A. De S. Maximal oxygen uptake, total metabolic energy expenditure, and energy cost in swimmers with physical disabilities. *International Journal of Performance Analysis in Sport* Volume 19, 2019. Issue 4. Pages 503–516
43. Kang M.-Y., Sapoval B. Prediction of maximal oxygen uptake at high altitude. *European Respiratory Journal* 2016 48: PA1583; DOI: 10.1183/13993003.congress-2016.PA1583
44. Bernardi L., Schneider A., Pomidori. L., Paolucci E., Cogo A. Hypoxic ventilatory response in successful extreme altitude climbers. *European Respiratory Journal* 2006 27: 165-171; DOI: 10.1183/09031936.06.00015805

45. Білошицький П.В., Онопчук Ю.М., Аралова Н.І. Математичні методи дослідження проблеми надійності функціонування організму за екстремальних умов високогір'я. *Фізіол.журнал.* 2003. Т. 49. № 3. С. 47–54.
46. Аралова А.А, Аралова Н.И., Белошицкий П.В., Онопчук Ю.Н. Автоматизированная информационная система функциональной диагностики альпинистов. *Спортивная медицина.* 2008. №1. С. 163–169
47. Колчинская. А.З., Моногаров В.Д., Радзиевский П.А., Молчанова Н.И. Комплексный контроль в альпинизме и его роль в управлении тренировочным процессом. *Управление процессом а-даптации организма спорт-сменов высокой квалифика-ции*: Сб.научн. трудов/ Ред. коллегия: Д.А. Полищук и др.- Киев: КГИФК, 1992. С.122–132
48. Колчинская А.З., Моногаров В.Д., Радзиевский П.А. Научно-методическое обепечение підготовки советских альпинистов к траверсу массива Канченджанги. *Теория и практика физической культуры*. 1991. 4. 10–14.
49. Колчинская А.З., Белошицкий П.В., Моногаров В.Д., Пвнутель Р.В., Радзиевский П.А., Красюк А.Н., Ивашкевич А.А., Борисов А.Н. Физиологическая работоспособность альпинистов в условиях экстремально низкого pO_2 во вдыхаемом воздухе. *Физиологический журнал.* 1989. 35, 2. С. 68–74.
50. Колчинская А.З., Лауэр Н.В., Шкабара Е.А. О регулировании кислородных режимов организма. В кн. «Кислородный режим организма и его регулирование». Киев: Наукова думка; 1966. С. 349–556.
51. Klyuchko O.M., Aralova N.I., Aralova A.A. Electronic automated work places for biological investigations. *Biotechnologia Acta*, K, 2019. V.12. №2. P. 5–26. https://doi.org/10.15407 /biotech12/02/005
52. Аралова А.А., Аралова Н.И., Ковальчук-Химюк Л.А., Онопчук ЮН. Автоматизированная информационная система функциональной диагностики спортсменов. *Управляющие системы и машины.*2008; 3:.73–78.
53. Аралова Н.И.Математическая модель механизмов краткосрочной и среднесрочной адаптации функциональной системы дыхания лиц, работающих в экстремальных условиях. *Кибернетика и вычислительная техника.* 2015.182.15–25.

Отримано 19.01.2022

Аралова Н.І.[1], д-р. техн. наук, старш .науч.співроб.,
старш .науч.співроб. відділу оптимізації керованих процесів
ORCID 0000-0002-7246-2736, e-mail: aralova@ukr.net
Білощицький П.В.[2], д-р. мед.наук, проф.,
Професор біологічного факультету
ORCID 0000-0002-6058-3602, e-mail: bilosh827@ukr.net
Зубієта-Калейа Г.[3], д-р. мед. наук, проф.,
директор
ORCID: 0000-0002-4283-6514, e-mail: zubieta@altitudeclinic.com
Аралова А.А.[1], канд. фіз.-мат. наук,
наук. співроб. відд. методів дискретної оптимізації,
математичного моделювання та аналізу складних систем
ORCID 0000-0001-7282-2036, email: aaaralova@gmail.com
[1]Інститут кібернетики імені В.М. Глушкова НАН України,
40, пр. Акад. Глушкова., 03187, Київ, Україна
[2]Уманський державний педагогічний університет імені Павла Тичини,
2, вул.. Садова, 20300 Умань, Черкаська обл.., Україна
[3]Інститут високогірної мелицини та патології Ла Пас, Болівія

АВТОМАТИЗОВАНА ІНФОРМАЦІЙНА СИСТЕМА ДЛЯ ОЦІНЮВАННЯ ПРАЦЕЗДАТНОСТІ АЛЬПІНІСТІВ В УМОВАХ НИЗЬКОГО pO_2 У ВДИХУВАНОМУ ПОВІТРІ

Вступ. Наразі в результаті постійного зростання інтенсивності людської діяльності, несприятливого для довкілля, необхідності виконання роботи в різних екстремальних умовах значно посилюються фізичні, психічні та емоційні навантаження на організм людини, що призводять до виражених змін у функціональних системах. Тому актуальним є завдання вивчення адаптації організму людини до роботи в екстремальних умовах. Робота альпіністів є досить адекватною моделлю для вивчення комбінованих ефектів гіпобаричної гіпоксії та гіпоксії фізичного навантаження. Необхідність оброблення великих обсягів інформації зумовлює необхідність використання сучасних комп'ютерних технологій, що дає змогу в процесі підготовки альпіністів багаторазово, майже в режимі реального часу прискорювати оброблення даних і накопичувати їх для подальшого використання для визначення поточного стану та прогнозування регуляторних реакцій організму на зовнішні та внутрішні порушення.

Метою роботи є розроблення автоматизованої інформаційної системи функціональної діагностики з використанням моделі регулювання кисневих режимів організму та її практичне застосування для вивчення стану висококваліфікованих альпіністів.

Методи. Методи програмування, створення автоматизованої інформаційної системи та методи функціональної діагностики.

Результати. На основі моделі регулювання кисневих режимів організму побудовано автоматизовану інформаційну систему функціональної діагностики осіб, які перебувають в умовах екстремальних порушень. Наведено результати апробації запропонованого програмного забезпечення для дослідження групи висококваліфікованих альпіністів.

Висновки. Запропоноване програмне забезпечення дає змогу використовувати модель кисневих режимів організму в режимі реального часу, тобто, багаторазово прискорює оброблення даних, дає можливість централізовано збирати інформацію для її попереднього оброблення, зберігання та колективного використання, уможливлює порівнювняння основних параметрів, які характеризують функціональну дихальну систему під час природних спортивних занять та отримані під час ергометричного навантаження.

Ключові слова: *методи функціональної діагностики, стан альпіністів високої кваліфікації, математична модель регулювання кисневих режимів організму, адаптація людини до роботи в екстремальному середовищі, гіпоксібрична гіпоксія.*

DOI: https://doi.org/10.15407/kvt208.02.082

KATRAKAZAS P., Ph.D.,
Research Area Manager
https://orcid.org/0000-0001-7433-786X, e-mail: p.katrakazas@zelus.gr
SPAIS I., Ph.D.,
Senior Project Manager
Researcher ID: https://www.semanticscholar.org/author/I.-Spais/1885927,
e-mail: ilias.spais@zelus.gr
Zelus P.C.,
Tatoiou 92, 14452, Metamorfosi, Athens, GR

BLUEPRINTS ELICITATION FRAMEWORK FOR AN OPEN ACCESS PAN-EUROPEAN NEURO-IMAGING ONLINE CENTRE

Introduction. *Recent infrastructural endeavours in the field of neuroscience aimed at data integration and sharing and availability of research output. This approach recognized that opening experimental results produces significant gains for science advancement. Nonetheless, this leaves a large part of the grassroots neuroscience community underutilized: access to neuroimaging infrastructures remains locally restricted, obstructing data acquisition and the means to investigate novel hypotheses.*

Purpose. *Within our paper we seek to address this gap by providing the blueprints for a delocalized e-neuroscience centre, opening the access to functional neuroimaging acquisition systems at a pan-European level. This aim will be achieved by building operational interoperability, standardizing, and integrating the services of neuroscience centres across Europe and the development of a virtual environment allowing all European researchers to acquire state-of-the-art neuroimaging data, exploiting the principles of the European Charter for Access to Research Infrastructures*

Results. *The implementation of all necessary actions for the harmonization and interoperability of the experimental procedures of the labs entail standardization of protocols, procedures in the form of consensus-based guidelines, harmonization of hardware and software set-up and availability across laboratories, as well as adopting of common standards and formats for acquired data and metadata structures.*

Conclusion. *Consistent and streamlined mobility processes aim to become a blueprint for networking of the overall neuroscience community. The harmonized process framework presented in this paper can facilitate better use from current and future neuroscience projects. Data economies of scale and recruitment streamlining will put local EU and international funds to better use than the now dispersed efforts. This will lead to more successful projects and better pacing for EU neuroscientific communities in the international stage.*

Keywords: *multi-centre interoperability, operational harmonisation, neuroimaging, sharing infrastructures, open access framework.*

INTRODUCTION

On the side of the COVID 19 pandemic, another, a silent pandemic is harnessing Europe in the last decades: the effects of brain disorders. Europe's leading cause of disability-adjusted life years (DALYs) is the effect of the central nervous system and brain disorders [1]. Yet, Europe has yet to develop a pan-European strategy that would respond effectively and decrease brain disorders' effect on reducing the quality of life. A horizontal approach cannot succeed due to the great diversity in brain disorders' manifestation and treatment across the different European regions [2]. Inter-regional differences in disorder manifestation that drive the corresponding inter-regional variability in brain disorder DALYs are unknown. The current infrastructural paradigm in functional neuroimaging, which grounds on the existence of a variety of distributed and independent centres, as well as equipped they are, cannot respond to the need: their experimental procedures and results are independent and incomparable. Only a pan-European neuroimaging centre that can run at an open access basis, the same procedures across different European Union (EU) regions simultaneously and generate validly harmonized results can identify the drivers of inter-regional variability in brain disorder manifestation.

PURPOSE

To develop detailed models of brain disorders' effects, incorporating data regarding inter-regional differences that can inform regional defined interventions and pan-European strategies which aim to reduce the disorder's systemic effects, the current research attempts to provide the rational towards a pan-European, decentralized, open and expandable infrastructure for transforming the ability to fight brain disease and understand brain health by incorporating inter-regional confounds of neurological disorder manifestation through a methodological and technological infrastructural harmonization. Grounded on the harmonization of existing infrastructures and processes our framework intends to exploit the notion that "the whole is greater than the sum of its parts".

PROBLEM STATEMENT AND AREAS OF CONSIDERATION

Background on interregional variability in brain disorders state of the art (S.O.T.A.)

Brain disorders are one of the greatest health challenges. Around 165 million Europeans are living with a brain disorder. European healthcare budgets are burdened by €800 billion per year, and this is expected to increase further as Europe's population ages and becomes more susceptible to brain disorders. Importantly, brain disorders and their effects show great inter-regional variability: some level of commonality in the prevalence of certain disorders exists, but a significant diversity in their manifestation and treatment across European regions results in a great diversity of their effects, as reflected in the inter-regional variance of disability-adjusted life years (DALYs) [1]. The rate of people in the world who die or are disabled from Alzheimer's and other dementias (which show the greatest increase over the last 25 years) ranges from 1109 DALYs per 100000 in Italy to 409 in Cyprus [3].

As far as ease-of-access is concerned, neuroimaging infrastructures that are remotely interlinked offer the opportunity for ease of access services that cover very specific needs of the researchers that will use them. Telepresence infrastructures like a mixed reality teleconferencing component are not unheard of in the healthcare sector and have been implemented in surgery teleconsultations amongst other uses. Remote operation of medical devices has been implemented numerous times in the context of telemedicine and can be readily implemented within the scope of the proposed infrastructure, supporting specificity in experimental execution. Focusing on the rigor that is required in neuroscientific experiments, an experimental protocol verifier tool has and can be implemented, that would consider the specifics of the equipment at hand, correcting or rejecting remote research protocols that are misaligned or unfeasible in the context of the infrastructure. Additionally, data pre-processing services or full processing services are within the scope of this infrastructure's ease of use and research usefulness suites. Furthermore, educational resources in the form of educational videos, online manuals, and knowledge base portals are staples in the support suite of every neuroimaging device and are provided as standard by their vendors.

FRAMEWORK DESCRIPTION

The inter-regional diversification of the DALYs of brain disorders in the population has been recently documented, but this does not stand also for the different disorder features. Our framework intends to allow the documentation of the disorder manifestation characteristics that drive inter-regional variability. Infrastructure modalities are needed as potentially standardized and harmonized means of intervention. Without knowing the elements that drive the differential effects of brain disorders inter-regionally, Europe cannot build a stable strategy to reduce the disease burden in a changing society. This goal falls beyond the scope and resources of any single country and must be addressed collectively. Up until now, the lack of resources and infrastructure for defining the inter-regional characteristics of brain disorders manifestation, has hampered progress in understanding the disease burden and improving brain health. The infrastructure needed for such an endeavour is grounded upon better coordination of research among European countries and beyond, as well as between academia and industry (i.e., science diplomacy).

The characteristics of such infrastructure include a decentralized structure that integrates functional neuroimaging hardware, processes, and analyses in such a way that interregional variation of neurological disorders can be studied directly at a pan-European level. Integrating different neuroimaging modalities, such as MEG, EEG, fMRI, and complemented by brain activity modulation techniques such as TMS and tDCS, this kind of structure will be able to achieve the full potential of methodological complementarity. Hence, a structure with procedural and technological infrastructural harmonization that operates inter-regionally and can be utilized by the European research community, to produce comparable results. This will then allow answering the scientific challenge of modelling the manifestation characteristics that drive the inter-regional diversity of DALYs.

The technical composition of such an infrastructure can be rather specific given the neuroscientific equipment (e.g., MEG, EEG, MRI, TMS, tDCS amongst others) involved. However, the specificity of data and equipment integration (e.g., stimulus synchronized MEG-EEG recordings, post tDCS EEGs, etc.) is always defined by the neuroscientific question. In that context, the case-by-case needs analysis is the framework that informs the specific infrastructural modalities that will be utilized in each case as well as their data integration and processing methodology. A questionnaire with closed answers can be created and disseminated among European scientists focusing on interregional variabilities of brain disorders. Such a questionnaire can be disseminated from social media, conferences, and networks of scientists/stakeholders in which the partners participate, reaching a wider acceptance and including questions that will define typical experimental paradigms and neuroimaging modalities needed by the community. The data gathered can then be analysed using descriptive statistics to ground the decisions of formatting our framework's structure.

While the ease of access provisions is not possible, nor desirable to be integrated into the proposed concept infrastructure, our framework aims to engage in a comprehensive participatory approach to involve the neuroscientific community in the process of selecting the most impactful and appropriate features for ease of access. A series of co-creation sessions will touch on every aspect described above. Thematic distribution in different sessions will ensure that clarity of purpose will be maintained. Agile principles will be transferred from IT to this needs analysis to streamline the participatory process. Engaging the full range of stakeholder target groups, user stories, personas and other established workflows of participatory UI design will be used in the co-creation sessions that will take place. The transnational dimension of stakeholder engagement will be maintained through the heavy use of teleconferencing and collaborative tools like interactive storyboards (e.g., Miro board). This process will identify the best impact/cost ease of access features that will empower the concept infrastructure.

Generally, in the context of every neuroscience research case there are already defined constitute parts that define and designate it as such. These include the core research question, which serves as the defining anchor of the research use case. Defined by this, but equally constant as a constituent part, is the equipment and methodology of the experimental setup for the use case. Equally important in the definition of every neuroscientific research use case are the necessary pre-processing and algorithmic processing approaches and the interplay of them with the data integration that is required from the relevant infrastructure. Also, as in every use case, experimental cohort composition as well as inclusion and exclusion criteria are also important constituent parts of every rigorous neuroscientific research use case. These use case features cannot, however be arbitrarily assessed against the overall scientific challenge of interregional differences of brain disorders.

Our framework intends to become more than a technical infrastructure for researching interregional brain disorder differences. As such, in the framework's recommendations for upscaling of the infrastructure there is the provision of identifying the field's major research axes. Through that endeavour our framework intends to formulate an authoritative, continuously updated knowledge base that will be able to identify the relevance of neuroscientific research cases within the specific field.

AN INNOVATIVE INTERPRETATIONAL MODEL

Multilayer modelling of interactions between brain functionality, cognition/behaviour, quality of life/dalys, and regional characteristics of health systems

The neuroimaging, behavioural, quality of life, and societal data can subsequently be incorporated in a multilayer network model, with brain functionality, cognition/behaviour, quality of life/DALYs, and regional characteristics of health systems as different layers. Taking advantage of the latest advancements in complex network science that led to the extension of graph analysis to the modelling of multilayer representations consisting of distinct networks [4], this network model will allow the detailed quantification of the interactions amongst those layers that drive inter-regional differences. This allows the modelling of global functional connectivity characteristics between the layers of the corresponding hypernetwork. As such, it has the potential to evolve our current two-dimensional understanding of the link between neurobiological deficits and DALYs into a multi-layered construct that effectively interprets the complexity of this relationship.

The challenge of merging connectivity metrics amongst the different disciplines will be overcome by estimating hyperedges that depict the interaction between nodes of the functionally distinct network layers in the graph [5]. Each measurement (i.e., brain functionality, cognition/behaviour, quality of life/DALYs, and regional characteristics of health systems) will constitute a distinct layer. Within-layer nodes depict the manifestation of the disorder in the corresponding layer, and their interaction pattern is modelled by a connectivity analysis that utilizes metrics dictated by the consequent discipline. Normalization will be applied to homogenize the networks across all layers and render them comparable to each other. Between-layer co-occurrence of the deficits will model the node-to-node interaction amongst different domains: The profiling of the co-occurring problems will be carried out using Spearman's rho and its output will attribute each node pair with a weight based on the co-occurrence of the corresponding symptoms. The corresponding hypernetwork of each individual will then be subjected to a statistical evaluation based on general linear modelling, like the one shown in [6]. Global network characteristics (e.g., resilience or topological characteristics such as clustering persistence diagram), appropriate for multilayer networks can also be estimated as indicators of the overall structure of the network. The structure of the model, having free weights on each node will be proposed as a template interpretational tool for later studies aiming to understand how other brain disorders affect the corresponding interregional DALYs variability.

Tackling global health challenges such as dementia has been directed toward the development of affordable, scalable, and broadly available biomarkers of brain connectivity [7]. Over the last two decades, M/EEG studies have benefited from the increased sophistication of processing pipelines, allowing researchers to gain more insight into the dynamics and connectivity of the brain [8]. This is illustrated by recent M/EEG studies incorporating graph theory, nonlinear dynamics, decoding, and whole-brain modelling, which bring novel opportunities for the study of dementia [9], [10].

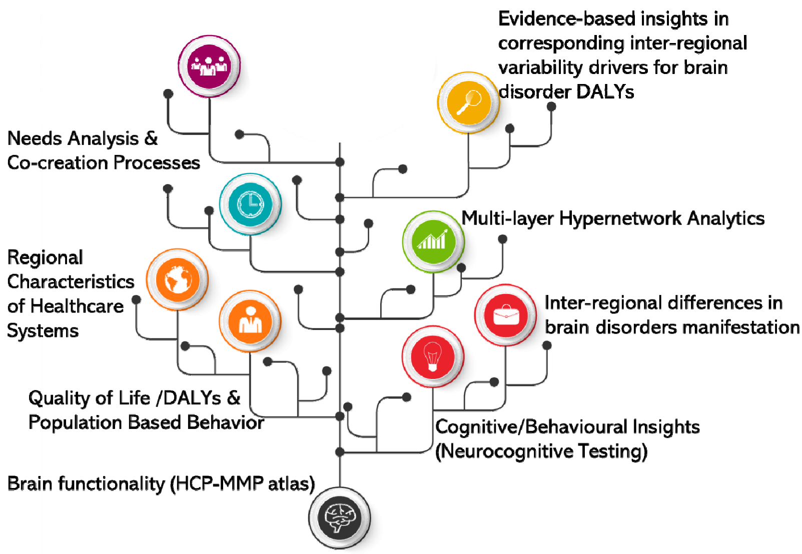


Fig. 1. Bird's Eye View of the Proposed Framework

The proposed approach based on the analysis of large volumes of standardized and harmonized data is expected to increase the efficacy of diagnostic tools and intervention strategies. Though the current and potential impact of functional connectivity studies based on M/EEG cannot be ignored, developing M/EEG multicentric studies is not an easy road and involves major technical difficulties and organizational needs. For instance, in a typical multicentric approach to task-related M/EEG functional connectivity, the major drawbacks arise from the task design, which needs careful validation. Furthermore, typical multicentric studies involving the acquisition of task-related M/EEG face several technical challenges, which include compensations for different stimulation timing and M/EEG synchronization procedures.

The framework therefore will consist of: (a) Acquisition protocol & Environmental settings and data quality measures, (b) Data sharing procedures, (c) basic pre-processing, (d) advanced pre-processing, and (e) a multilayer framework for M/EEG connectivity. Standardization and harmonization applied within the framework workflow will include: (i) a harmonization of M/EEG measurements to attenuate any variability linked to lightning, acoustic background noise, electromagnetic noise of the M/EEG chamber and other setting conditions, but also between resting-state (eyes-open and eyes-closed) measurements i.e., room lighting, and ERP measurements i.e., compensation of time delays, synchronization procedures, (ii) a standardization for data sharing i.e., following organizational standards for data sharing and applying an annotation and classification system, (iii) a standardization of basic and advanced pre-processing steps such as re-referencing, filtering, rejecting bad channels and bad segments, artifact rejection, etc, leading to spatial harmonization, data normalization of data from healthy participants as well as patient-control normalization and (iv) the estimation of a multilayer network of each patient phenotype.

An important step in M/EEG harmonization will be the measure of the harmonization success an essential element for guiding further data analyses and planning future preclinical studies. Such measures will consider the acquisition system and relevant acquisition parameters as covariates in the framework M/EEG studies of connectivity. Acquisition systems can make a significant contribution to the variance of multisite M/EEG studies, particularly when they differ in the type of electrodes/sensors, and electronics. Metrics indicating the success of the harmonization will refer to source space and will help to build methodological consensus, which in turn will contribute to establishing functional connectivity measures as reliable biomarkers for the early identification of subtypes within a particular neurodegenerative disorder and facilitate intervention. This early identification system implies that the traditional classification of neurodegenerative disorder based on group analyses shall be complemented with individual-level analyses. Finally, the methodology for sharing data and its ready-to-use processing pipelines will be critical for the replicability and cross-validation of the framework studies.

Development of a horizontal interconnection of the labs in terms of infrastructure operation. Currently, the operation of each infrastructure is executed via direct commands on the manufacturer's software defining the corresponding measurement parameter. The different software uses similar but not identical names for each command while the user interface is diversified in accordance to the manufacturer's style. This ensures operability from each lab's technician or operator but obstructs the option of one operator knowing and acting at several systems. The suggested framework intends to ground its horizontal interconnection on a system consisting of (a) a Sandboxed Data Container, (b) an authentication system, (c) an SDK handling communication with different manufacturers' operating system of each infrastructure and (d) a General User Interface (GUI) operated digital twin of a simplified version of manufacturer's software. Specifically, it will develop digital twins of each software, running at a Sandboxed Data Container (SDC), ensuring safety of the system. The digital twin of each software will have a simplified General User Interface, and commands that are to be utilized for fine-tuning the parameters of experimental paradigms. This deadlock is one of the core aspects that SDC tries to overcome through by-design mechanisms where the data, processes, metadata, and the intermediate results are safeguarded and protected by any unauthorized entity (Event Handler tier). Software commands which are not to be handled by external operators (for system's stability and safety reasons) will not be visible and will remain functional only in the original software of the system manufacturer. These features can cover the aspects of privacy and security, so the SDC can ensure to the data owners that their data is secured and monitored (Monitoring tier) in a way that no one else can intervene. The names of the commands that are to be used will be unified across the labs.

In more detail, the instantiated SDC includes the researcher's credentials, which are used by an Authentication Server (AS) to grant the access. Hence, when an external operator of the virtual infrastructure (after receiving relevant permission and login info from the common user interaction gate) enters the system as operator, s/he will see a unified user interface independently of the system, with common names on each parameter to be operated. Inside the SDC realm, each data is handled as a unique entity and is governed (Data Service & Storage tier) with different constraints and regulations. This is the case of interest for many scenarios since each person must have the freedom to define personalized policies of his/her own data. This approach will allow experienced experimenters (after a short virtual educational seminar on operation of the unified system) to be able to handle all infrastructure operation that will be used for recording of the corresponding experimental paradigm, while maintaining system's stability. The corresponding commands will be sent to the original software to run the analogous functionality. One key principle that the system will cover is the adherence to the different policies used by each hospital or institute on data management rights. Important points of this approach are (a) the sandbox's operation, (b) the homogeneity of GUI of the digital twins of all different software, (c) the choice of the parameters and operation commands that will be available to external – virtual users. The later will have to be defined in relation to decisions made based on the experimental paradigms that will be offered, as well as on the basis of system's stability.

Development of a vertical interconnection of the labs and common user interaction gate in terms of infrastructure operation. In line with the overall objective of unifying data and processes from individual laboratories in the neuroimaging domain, the framework should be supported technically by a scalable infrastructure and a set of data handling tools. The infrastructure will receive data from the individual laboratories and will apply data quality analysis, anonymisation and transformation to provide them to the researchers upon relevant requests. Moreover, a two-way communication link between the researchers and the labs will be established to allow conducting experiments based on specific protocols. The architectural overview of the technical infrastructure is shown in the Figure 2.

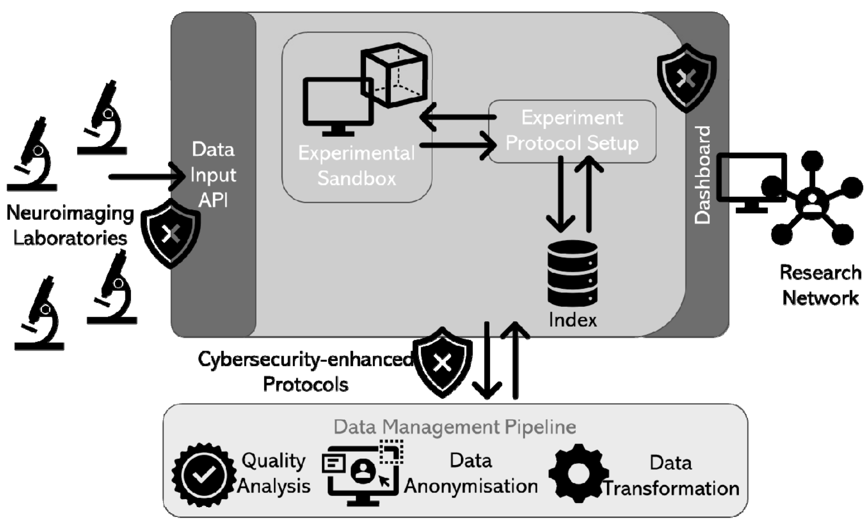


Fig. 2. Architectural Overview of the Framework's Technical Infrastructure

The infrastructure comprises a data input API built in REST so that it allows implementing different connectors for the data and experiments of each laboratory. The input will be inserted based on an event streaming platform, such as Apache Kafka, to ensure high throughput and high availability. The system will also involve a data management pipeline so that the data are initially analyzed and anonymized, ensuring that any sensitive identifiable information is not forwarded to the user. Furthermore, data transformation will be supported in order to produce an output format that will be suitable for data presentation/visualization to the researchers.

The core component of the infrastructure is the communication layer, which will be used to conduct and retrieve experiments performed in different laboratories. From the perspective of the researcher, an experiment dashboard will be used as a unified environment that allows exploring the data and equipment availability of the different labs. The researcher will be able to upload a new experiment request via an intuitive UI, using an appropriate protocol for the experiment. All experiments, along with the relevant protocols will be stored in a searchable index, allowing easy retrieval. This way it will be possible to check whether new experiments are compliant with existing protocols and/or provide recommendations based on AI to make them compliant.

Upon determining a relevant lab, in terms of data and availability, the researcher will be able to run the experiment in a virtual sandbox. The sandbox is built using virtualization technologies (e.g., Docker), ensuring that the laboratory data are secure from potential compromises, either intentional or not. State-of-the-art security protocols will be used to connect to virtual instances of the sandbox. The data produced by the submitted experiment will be forwarded to the data management pipeline. The pipeline will initially perform quality checks, focusing on curating (pre-processing is performed in the labs). After that, the data will be anonymized using state-of-the-art anonymization techniques to ensure that all valuable information is maintained while removing any personally identifiable information. The final step will be that of data transformation. The researcher will have control over the format of the data and will be able to request for different statistics and/or visualizations. The results of the system will be presented to the users in a dashboard, while they will also be available using a RESTful API. Therefore, several visualizations will be available to aid the interacting parties in understanding the output and providing useful suggestions. Focus will be given on statistics and data aggregation techniques, while full data retrieving as streams will also be supported.

Secure data movement techniques. Protecting data, either in transit or at rest, is a necessity that plays a catalytic role, especially in cloud computing infrastructures. The most basic countermeasure that ensures data privacy are properly configured access control policies. Discretionary Access Control (DAC), Mandatory Access Control (MAC), Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) are some approaches. However, these do not provide absolute security, and sensitive data might still get exposed in case of attacks, especially during the data transmission between source and destination (e.g., man-in-the-middle attacks). In all cases, data encryption by the data owner before it is uploaded to the cloud is necessary; in this way confidentiality is guaranteed even in case of breaches and leaks.

However, classic encryption techniques, such as Advanced Encryption Standard (AES), Data Encryption Standard (DES), and Triple Advanced Encryption Standard (3DES), are not feasible in cloud computing settings due to their complexity and lack of flexibility, scalability, and fine-grained access control [11], [12]. Instead, cloud-suitable alternatives are Attribute Based Encryption (ABE) and Identity Based Encryption (IBE). Traditional encryption mechanisms involve the data owner encrypting the data with a user's public key before uploading them. The user can then decrypt the data using their private key. In IBE, public keys are calculated based on a unique identifier of a user, such us their e-mail address. This eliminates the need of public key retrieval [13]. The private key is distributed by a private key generator. ABE (also known as fuzzy identity encryption) uses a combination user attributes instead, such as their name, city, or street number to generate public/private keys [14]. In addition to data confidentiality, data integrity is another critical element in data security. Data integrity refers to protecting data from unwanted modification, deletion, or fabrication. The first step in ensuring it, is proper access control through authorization. In addition, RAID strategies store data in different locations to provide redundancy and fault tolerance [15].

Finally, Data Integrity Validation (DIV) techniques provide assurance to users that data is intact and has not undergone corruption, deletion or modification. RSA and MD5 are some of the simplest techniques, where downloaded data signatures are calculated and compared to a reference. In addition to these, public auditing techniques offer a scalable alternative, where a Third-Party Auditor (TPA) is introduced to handle the interaction with the Cloud Storage Server [16]. Finally, where preventive measures fail, timely detection of the failure is necessary. Intrusion Detection Systems (IDS) monitor an infrastructure at the network or file-system level for abnormal behaviour as well as unauthorized data access [17]. These, in turn, can trigger appropriate mitigation actions, such as Moving Target Defenses (MTD) which can significantly increase an attacker's uncertainty and decrease their chance of success [18].

Co-creative iteration on all components. Co-creation as a part of the field of participatory design has received significant attention in the context of innovation in recent years. Due to the efficiency and adaptability in responding to diverse and changing environments and contexts, this has been identified as a potential booster for the implementation of innovative and experimental solutions [19]. The main point of co-creation is the transformation of passive actors such as end-users into active ones. Hence, they are actively involved in the development processes of products, services and systems [20] as well as the definition and creation of common values taking all actors and their needs into account [21]. Users and actors are not considered just during research process in co-creation methodology, but also throughout the production phases such as inspiration, co-design, testing, and execution of a solution. This includes the co-production process [22]. From a business point of view, this active participation in co-creation processes usually aims to shift the focus from a business centric one towards personalized and satisfying customer experiences [23]. These characteristics of co-creation methodology and its value led to expanding the fields of application as well as the notion of concept. More specifically, it has been attempted as a promising means to engage neglected actors and stakeholders in various domains of innovation (e.g., in public sector innovation) and as a way to set up collaborative processes like those that are needed to better involve in innovation [24].

Our framework intends to guide pilots which should be accommodated by the infrastructure-driven partners, intrinsically adopting the co-creation methodology as its principal objective. This refers to a systematic user co-creation approach integrating research and innovation processes in real life communities and settings. In practice, the exploitation of qualitative and quantitative data of the needs analysis can shape the plan of integration and harmonization of the infrastructure and services as well as the configuration of the blueprint that will depict the vertical infrastructure, subserving the common gate of experimental protocol and user interaction. In the following, those plans will be forwarded via a second iteration of questionnaire dissemination and feedback sessions addressed to a shorter list of scientists of the ERA that focus on inter-regional brain differences and relevant targeted stakeholders. As such this process will employ co-creative elements in an iterative feedback process that will align the concept of the infrastructure and its implementational details.

Blueprints for Infrastructure scaling up and processes generalization
Technical Scale-up Blueprint. Current infrastructural paradigm of functional neuroimaging relies rely on various distributed and independent centres, that cannot serve the necessity for experimental procedures and results to be relevant and comparable. An increasingly high cost of acquiring and maintaining a neuroimaging facility offering all modalities, restricts the availability of such resources at a European level. Hence, the local access to neuroscientific infrastructures offering high-quality neuroimaging services for all imaging modalities provided by the framework is currently uncommon across the European area, restricting the full use of the operational time of the corresponding infrastructures (several such infrastructures currently have free operational time, as their use is based on experimenters performed from a small number of researchers), and the exploitation of the scientific knowledge and expertise of researchers which do not have access to such infrastructure. Therefore, the framework's goal is to create the ground for a modular system that will be able not only to integrate the different infrastructure horizontally, but also to support the whole variability of experimental needs of the ERA scientists focusing on inter-regional variability of brain disorders. Hence, a modular system that will be utilizable for different experimental use cases. Thus, a set of operations and services that are needed to answer the community's needs will be defined and a modular system will be described in detail that will be able to implement the different experimental procedures across the different labs. Hence, a system that will not only include in the sandboxed operated digital twin, the software commands that are needed to run the resting state analysis of defining Subjective Cognitive Decline manifestation, but also all the software commands that are needed to run all the experiments defined by the community via the needs analysis. Technical requirements that are needed for a new infrastructure's operating system to integrate to the framework system and to connect to the corresponding SDK will be defined and published accordingly. Scientific and methodological requirements regarding the services and the lab's support for the system to run fluently will be defined, comprising of best practices for each neuroimaging facility.

Alongside, a common technical framework for compiling novel use case scenarios, novel experimental paradigms and research protocols will be defined.

This will include global definition of terms, so that a common set of procedures is easy to define across labs, and a definition of neuroimaging modalities used under the umbrella of the framework canvas so that that the protocol is easily evaluated on the basis of the functionality of the integrated system. A template protocol allowing specific definitions on the parameters (the ones applicable on the integrated framework system of the infrastructure) will be prepared while initially, an algorithm will evaluate applicability of the protocol in the integrated system. At a second stage a technician/researcher of, familiar with the functionality of the integrated framework system will evaluate further details. As soon as this two-stage procedure is completed, and the common technical framework of a newly submitted experimental protocol is respected, the research will be prioritized for implementation. These elements will comprise the technical blueprint of upscaling the number and variability of infrastructure included, allowing integration of other European labs in the common functionality of the virtual decentralized system.

Scientific Diversification Blueprint. Several brain imaging projects have attempted to identify suitable biomarkers of neurodegenerative disease. For instance, the Alzheimer's Disease Neuroimaging Initiative (ADNI) contributed to the development of blood and imaging biomarkers, the understanding of the biology and pathology of aging, and to date has resulted in over 1800 publications. ADNI also impacted worldwide ADNI-like programs in many countries worldwide. Though, these previous mega-studies have contributed to the discovery of potential mechanisms and biomarkers of multiple brain disorders, most of these imaging biomarkers have a relatively small effect sizes and the study results were drawn from multi-site data which are often heterogenous and used now outdated traditional low-resolution data acquisition protocols. In addition, there have been no human brain M/EEG studies that explore multiple neurological disorders that occur through the lifespan within the same cohort of subjects. Currently, typical multi-site studies for studying brain disorders have great advantages such as: (i) quicker recruitment of the necessary number of subjects, (ii) clearer results as the patient sample of multi-site trials is supposed to be representative and can potentially reach for more generalizable findings, (iii) sharing resources amongst collaborative sites, and (iv) promoting networking. However, multi-site trials require strong efforts for quality assurance concerning admission, treatment and follow-up, thus a highly developed coordinating centre is needed. In addition, typical multi-site neuroimaging studies have been based on a single or small number of the same model M/EEG systems at different sites and thus did not fully address standardization of the data acquisition across different scanner models or vendors. In addition, such studies usually require the procurement and/or distribution of common neuroimaging (e.g., portable EEG systems) devices of same specifications to attenuate the variability in the expected results. However, this practice can significantly increase the cost of such studies, while it can introduce errors even in the acquisition phase as the researchers of a centre may not be entirely familiar with the newly operated devices. In line with this, this practice likely excludes the use of MEG systems if sites haven't procured in the past as the systems require a dedicated space for installation, and cost from €1.5–4 million.

Our framework aims to accelerate harmonization technologies to be used in the integrated neuroimaging centres by combining approaches to high-quality imaging acquisition, pre-processing, study design, and statistical bias correction to potentially improve the sensitivity and validity of imaging results. The conceptual framework for major health challenges will be based on the notion that as patients are not the same, regions are not the same, thus it will follow a population centred approach and take into account differences of the EU society compared to the available literature, metanalyses, clinical trials and guidelines. These observed differences will improve understanding of disease pathophysiology and help researchers to tailor their approach and research questions based on the population characteristics. All use case scenarios that address global inter-regional health challenges will be introduced into the relevant ecosystem via a common scientific framework. This will include forming hypotheses regarding the Impact of inter-regional differences in each health challenges and thus forming appropriate research questions that can account for account inter-regional differences. A unified template protocol will set the use cases' endpoints which will include features and measures derived from the imaging modalities and will provide a benchmark against which other phenotypical data will be assessed. These endpoints will provide objective evidence of beneficial and adverse outcomes.

Though aims will differ across the use cases, the framework will drive the main endpoints but also serve as an exploratory tool that contributes to the understanding of disease mechanisms. Studies under the framework's umbrella are a response to the increasing demand for more data, but can also promote collaborations across institutions and countries, and experts worldwide will have access to the large data sets and can combine their group expertise. Thus, a framework for new data syntheses will be defined, ensuring that appropriate technologies and strategies are used to manage the large amounts of data generated.

Policy upscaling Blueprint (Developing a unified financial, administrative, and regulatory framework exemplar for EU multicentric Research and Innovation). The unified financial, administrative, and regulatory framework exemplar for EU multicentric Research and Innovation pertains to regulatory issues and problems related to Copyright and Data Protection in the field of Research and Innovation which may be resolved in a way that allows for successful administrative and financial planning. Regarding Copyright law regulatory issues, we assess that the proposed framework's successful operation depends on the exploitation through the Internet of information goods and/or databases of information goods which will be distributed Europe-wide. The excludability imposed by Copyright's legal edifice to said information goods and databases is a regulatory issue to cope with through Openness in Copyright law and the new mandatory exceptions and limitations introduced through Directive 2019/790/EU.

In addition to Copyright law issues, planning entails the successful handling of data protection issues. The envisioned framework should have to cope with proper action for the implementation of the General Data Protection Regulation's mandates, i.e., Regulation 2016/679/EU. Therefore, personal data protection considerations will also have to be assessed carefully. The GDPR must be considered in relation to the suggested setup and operation from a viewpoint that meets the requirements of the adoption of privacy-by-design and advanced security techniques. GDPR compliance is a task related to internet security of the infrastructure of it, thus compliance with the requirements of Directive 2016/1148/EU concerning measures for a high common level of security of network and information systems across the Union as well as the acquisition of proper European Cybersecurity Certification provisioned in Regulation 2019/881/EU could be considered.

Most likely ethical issues will be faced, too, at least in the sense of drawing up and/or imposing a code of conduct for scientific research and innovation. For example, ethical restrictions imposed by Regulation 2014/536/EU and the Principles of European Medical Ethics adopted by the European Council of Medical Orders on 06/01/1987 including the principles adopted on 06/02/1995 or the ethical principles adopted in Kos Island (Greece) of the European Charter of Medical Ethics in 10/06/2011 will have to be considered.

In addition to the regulatory framework as outlined above, the new infrastructure should define prioritization algorithms based on excellence for admitting new protocols/studies from all ERA scientists, working as an open access infrastructure. The algorithm will define in detail, which research protocols are undertaken by the new infrastructure, which are prioritized over other in case that more than one is submitted, how the work is distributed across the partners, and under which regulations performing the study will be realized.

CONCLUSION

The current paper presents a pan-European functional neuroimaging framework, that will be able to model inter-regional differences in brain disorder manifestation. The suggested framework will rely on the principles of open science, offering its services and high-end decentralized infrastructures via virtual access to all European scientists based on an "access through excellence" gate. Expanding on the opening of science via the first (i.e., open access papers) and second (open data) pillar, the pan-European centre will open access to the third pillar of science: infrastructure. It will include the means and procedures to perform state-of- the-art experimental protocols proposed by the community of European neuroscientists targeting inter-regional differences for a variety of brain disorders.

In order for the framework to be developed it will require a horizontal harmonization of infrastructures and standardization of procedures by a decentralized set of state-of-the-art laboratories guaranteeing validity and comparability of research results obtained from a variety of labs participating across different European regions. It should also have a unique and common virtual access gate for all integrated labs and all European scientists aiming to utilize its services offering a simple, yet a user-friendly way for each scientist to submit a research protocol that will be able to run inter-regionally. Thus, it will offer its services via remote access incorporating digital twins of all systems in the cloud, and thereby allow the running of experiments simultaneously in the different labs, without the physical presence of the researcher in all of them.

The suggested framework is designed so as to offer the possibility to every neuroscience centre across Europe to participate as a hub of the pan-European structure following guidelines and procedures of integration. It defines a novel organizational model operating as a unique but decentralized structure of integrated working hubs, able to pool and share resources in studying inter-regional differences in brain disorders manifestation. Built upon a set of "small-world" network properties, the framework allows the functional integration of highly specialized research hubs into a working virtual neuroscience over-structure.

Finally, we present the blueprints of a unified financial, administrative and regulatory framework exemplar for EU multicentric Research and Innovation pertains to the regulatory issue. Following ESFRI principles and drawing upon similar structures such as the Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS), the framewok defines procedures to accept the implementation of research protocols submitted openly by the community. Moreover, it defines the prioritization of studies submitted, means of financial compensation of the hubs for executing the research and accountability. The opening of infrastructures and procedures via pre-defined principles will answer the risk of fragmentation of funding after the end of some of the large-scale programs in 2023 and systematically align the research agendas to a synergistic commitment. Overall, it will define actions related to copyright law issues, data protection, cybersecurity, ethical and contract/business issues as a pan-European structure. The blueprint will also define ways of collaboration, research refinement, and complementarity with other major initiatives in the EU (EBRI, EBRAINS, the JPND joint program, the Human Brain Project) and across (Brain Initiative) that will allow means of exchange and implementation of interregional research services.

REFERENCES

1. Ravindranath V. *et al.* Regional research priorities in brain and nervous system disorders. *Nature*, vol. 527, no. 7578, Art. no. 7578, Nov. 2015, doi: 10.1038/nature16036.
2. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol.*, vol. 16, no. 11, pp. 877–897, Nov. 2017, doi: 10.1016/S1474-4422(17)30299-5.
3. Vos T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, vol. 386, no. 9995, pp. 743–800, Aug. 2015, doi: 10.1016/S0140-6736(15)60692-4.
4. De Domenico M., Granell C., Porter M.A., and Arenas A. The physics of spreading processes in multilayer networks. *Nat. Phys.*, vol. 12, no. 10, pp. 901–906, Oct. 2016, doi: 10.1038/nphys3865.
5. Battiston F., Nicosia V., Latora V. The new challenges of multiplex networks: Measures and models. *Eur. Phys. J. Spec. Top.*, vol. 226, no. 3, pp. 401–416, Feb. 2017, doi: 10.1140/epjst/e2016-60274-8.
6. Paraskevopoulos E., Chalas N., Anagnostopoulou A., Bamidis P.D. Interaction within and between cortical networks subserving multisensory learning and its reorganization due to musical expertise. *Sci. Rep.*, vol. 12, no. 1, Art. no. 1, May 2022, doi: 10.1038/s41598-022-12158-9.

7. Mantzavinos V. and Alexiou A. Biomarkers for Alzheimer's Disease Diagnosis. *Curr. Alzheimer Res.*, vol. 14, no. 11, pp. 1149–1154, 2017, doi: 10.2174/1567205014666170203125942.
8. Larson-Prior L.J. *et al.* Adding dynamics to the Human Connectome Project with MEG. *NeuroImage*, vol. 80, pp. 190–201, Oct. 2013, doi: 10.1016/j.neuroimage.2013.05.056.
9. Dottori M. *et al.* Towards affordable biomarkers of frontotemporal dementia: A classification study via network's information sharing. *Sci. Rep.*, vol. 7, no. 1, Art. no. 1, Jun. 2017, doi: 10.1038/s41598-017-04204-8.
10. Iakovidou N.D. Graph Theory at the Service of Electroencephalograms. *Brain Connect.*, vol. 7, no. 3, pp. 137–151, Apr. 2017, doi: 10.1089/brain.2016.0426.
11. Shabir M.Y., Iqbal A., Mahmood Z., and Ghafoor A. Analysis of classical encryption techniques in cloud computing. *Tsinghua Sci. Technol.*, vol. 21, no. 1, pp. 102–113, Feb. 2016, doi: 10.1109/TST.2016.7399287.
12. Lingwei S., Fang Y., Ru Z., and Xinxin N. Method of secure, scalable, and fine-grained data access control with efficient revocation in untrusted cloud. *J. China Univ. Posts Telecommun.*, vol. 22, no. 2, pp. 38–43, Apr. 2015, doi: 10.1016/S1005-8885(15)60637-9.
13. Lee K., Lee D.H., Park J.H .Efficient revocable identity-based encryption via subset difference methods. *Des. Codes Cryptogr.*, vol. 85, no. 1, pp. 39–76, Oct. 2017, doi: 10.1007/s10623-016-0287-3.
14. Pasupuleti S.K. and Varma D. Chapter 5 — Lightweight ciphertext-policy attribute-based encryption scheme for data privacy and security in cloud-assisted IoT. in *Real-Time Data Analytics for Large Scale Sensor Data*, vol. 6, H. Das, N. Dey, and V. Emilia Balas, Eds. Academic Press, 2020, pp. 97–114. doi: 10.1016/B978-0-12-818014-3.00005-X.
15. Warsinske J. *et al. The Official (ISC)2 Guide to the CISSP CBK Reference*. John Wiley & Sons, 2019.
16. Zhou L., Fu A., Yu S., Su M., and Kuang B., Data integrity verification of the outsourced big data in the cloud environment: A survey. *J. Netw. Comput. Appl.*, vol. 122, pp. 1–15, Nov. 2018, doi: 10.1016/j.jnca.2018.08.003.
17. Khraisat A., Gondal I., Vamplew P., and Kamruzzaman J. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, vol. 2, no. 1, p. 20, Jul. 2019, doi: 10.1186/s42400-019-0038-7.
18. Nor'a M.N.A. and Ismail A.W., Integrating Virtual Reality and Augmented Reality in a Collaborative User Interface. *Int. J. Innov. Comput.*, vol. 9, no. 2, Art. no. 2, Nov. 2019, doi: 10.11113/ijic.v9n2.242.
19. Sarasvuo S., Rindell A., and Kovalchuk M.. Toward a conceptual understanding of co-creation in branding. *J. Bus. Res.*, vol. 139, pp. 543–563, Feb. 2022, doi: 10.1016/j.jbusres.2021.09.051.
20. Saarijärvi H. The mechanisms of value co-creation. *J. Strateg. Mark.*, vol. 20, no. 5, pp. 381–391, Aug. 2012, doi: 10.1080/0965254X.2012.671339.
21. Rizzo F., Deserti A., and Komatsu T. Implementing social innovation in real contexts. *Int. J. Knowl.-Based Dev.*, vol. 11, no. 1, pp. 45–67, 2020.
22. Frow P., Nenonen S., Payne A., and Storbacka K. Managing Co-creation Design: A Strategic Approach to Innovation. *Br. J. Manag.*, vol. 26, no. 3, pp. 463–483, 2015, doi: 10.1111/1467-8551.12087.
23. Martínez-Cañas R., Ruiz-Palomino P., Linuesa-Langreo J., Blázquez-Resino J.J. Consumer Participation in Co-creation: An Enlightening Model of Causes and Effects Based on Ethical Values and Transcendent Motives. *Front. Psychol.*, vol. 7, 2016. Last accessed: May 21, 2022. URL: https://www.frontiersin.org/article/10.3389/fpsyg.2016.00793
24. Jansma S.R., Dijkstra A.M., and de Jong M.D.T. Co-creation in support of responsible research and innovation: an analysis of three stakeholder workshops on nanotechnology for health. *J. Responsible Innov.*, vol. 9, no. 1, pp. 28–48, Jan. 2022, doi: 10.1080/23299460.2021.1994195.

Received 23.05.2022

Катраказас П., Ph.D.
менеджер наукового напряму,
https://orcid.org/0000-0001-7433-786X, p.katrakazas@zelus.gr
Спаіс І., Ph.D.,
старший менеджер проекту,
ilias.spais@zelus.gr,
Researcher ID: https://www.semanticscholar.org/author/I.-Spais/1885927
Zelus P.C.,
Tatoiou 92, 14452, Metamorfosi, Athens, GR

ПЛАТФОРМА ВИЯВЛЕННЯ СТРУКТУРИ ДЛЯ ЗАГАЛЬНОЄВРОПЕЙСЬКОГО ОНЛАЙН-ЦЕНТРУ НЕЙРОВІЗУАЛІЗАЦІЇ З ВІДКРИТИМ ДОСТУПОМ

Вступ. Нещодавні інфраструктурні зусилля в галузі нейронаук було спрямовано на інтеграцію та обмін даними та доступність результатів досліджень. Цей підхід визнав, що відкриття експериментальних результатів дає значні переваги для прогресу науки. Тим не менш, це залишає значну частину масової нейронаукової спільноти недостатньо інформаційно забезпеченою: доступ до інфраструктури нейровізуалізації залишається локально обмеженим, що перешкоджає збиранню даних і використанню засобів для дослідження нових гіпотез.

Метою статті є прагнення усунути цю прогалину, надавши платформу для делокалізованого центру електронної нейронауки, що відкриває доступ до функціональних систем нейровізуалізації на загальноєвропейському рівні. Ця мета буде досягнута шляхом побудови оперативної сумісності, стандартизації та інтеграції послуг нейронаукових центрів по всій Європі та розроблення віртуального середовища, що дасть змогу всім європейським дослідникам отримувати найсучасніші дані нейровізуалізації, використовуючи принципи Європейської Хартії доступу до дослідницьких інфраструктур.

Результати. Реалізація комплексу необхідних дій для гармонізації та забезпечення сумісності експериментальних процедур різних лабораторій передбачає стандартизацію протоколів, процедур у формі настанов на основі консенсусу, гармонізацію налаштувань апаратного та програмного забезпечення та доступності в лабораторіях, а також прийняття спільних стандартів і форматів для отриманих даних і структур метаданих.

Висновок. Послідовні та впорядковані процеси мобільності мають за мету стати структурою для створення мережі спільноти нейронауковців у цілому. Узгоджена структура процесу, подана у такому документі, може сприяти кращому використанню поточних і майбутніх проектів нейронауки. Економія даних за рахунок масштабування створених програмних продуктів та оптимізація впорядкування надасть можливість використовувати місцеві кошти ЄС і міжнародні кошти краще, ніж нині у разі розпорошених зусиль. Це приведе до успішніших проектів і кращого темпу поповнення вагомими результатами нейронаукових спільнот ЄС на міжнародній арені.

Ключові слова: *багатоцентрова сумісність, оперативна гармонізація, нейровізуалізація, інфраструктури спільного використання, структура відкритого доступу.*

ДО УВАГИ АВТОРІВ!

У журналі надано результати досліджень у галузях теорії та практики інтелектуального керування, інформатики та інформаційних технологій, а також біологічної та медичної кібернетики.

Цільова аудиторія — науковці, інженери, аспіранти та студенти вищих навчальних закладів відповідного фаху.

Вимоги до рукописів статей

1. Рукопис надають на папері у двох примірниках (мова — англійська, українська, 17–22 с.) та електронну версію. До рукопису додають:

• анотації — українською та англійською мовами (прізвище, ініціали автора/ів, місце роботи, місто, країна, назва статті, текст 250–300 слів, з виділенням рубрик: вступ, мета, результати, висновки, ключові слова 5–8 слів);

• список літератури мовою оригіналу — у порядку згадування в тексті, за стандартом ДСТУ 8302:2015;

• список літератури — переклад джерел англійською мовою, прізвища та ініціали авторів — транслітерація;

• ліцензійний договір;

• відомості про автора/ів українською та англійською мовами повинні містити: ПІБ, вчений ступінь, наукове звання, посада, відділ, місце роботи, поштова адреса організації, телефон (для зв’язку редактора), авторські ідентифікатори ORCID або ResearcherID, E-mail.

2. Текст статті подають з обов'язковими рубриками: вступ, постановка завдання/проблеми, мета, результати, чітко сформульовані висновки.

Вимоги до текстового файлу

Формат файлу * .doc, * .rtf. Файл повинен бути підготовлений за допомоги текстового редактора Microsoft Word.

Використовувані стилі: шрифт Times New Roman, 12 пт, міжрядковий інтервал – 1,5. Формат паперу А4, всі береги — 2 см.

Формули набирають у редакторах формул Microsoft Equation Editor 3.0. та MathType 6.9b. Опції редактора формул — (10,5; 8,5; 7,5; 14; 10). **Ширина формул — до 12 см**.

Рисунки повинні бути якісними, створені вбудованим редактором рисунків Word Picture або іншими Windows-додатками (рисунки надають окремими файлами відповідних форматів). **Ширина рисунків — до 12 см**.

Таблиці виконують стандартним вбудованим у Word інструментарієм «Таблиця». **Ширина таблиці — до 12 см**.

Передплату на журнал (друкована версія) в Україні здійснюють:

- за «Каталогом видань України», індекс передплати друкованої версії — 86598;
- за допомоги передплатної агенції «Укрінформнаука» НАН України, ukrinformnauka@gmail.com, індекс журналу — 10029