Forming the Architecture of a Multi-Layered Model of Physical Data Storage for Complex Telemedicine Systems

Dmitry V. Polezhaev [1], Aslan A. Tatarkanov [1*], Islam A. Alexandrov [1]

[1] Institute of Design and Technology Informatics, Russian Academy of Sciences, Moscow, Russian Federation.

Abstract

The relevance of this research is determined by the need to study the issues of improving data storage technologies for complex telemedicine systems. The objective is to create a multi-layered data storage model for complex telemedicine systems to ensure the most complete use of their capacity and the timely expansion of existing storage. The research is conducted on the basis of an analysis of existing opportunities and problems in the field of data storage technologies. An analysis of the main features of the development of data storage technologies revealed that the existing models have no detailed description of the recording and physical storage of data bits, which is necessary for describing the storage process. Different architectures are reviewed, and their strengths and weaknesses are discussed. Within the framework of a demonstration experiment using the Kohonen neural network apparatus as a tool for solving the problem of placing objects in accordance with the required parameters, it is shown that the proposed storage system resource management model is operable and allows solving the problem of rational use of physical resources. As a result, a multilevel model of data storage is proposed, which combines the levels of storage process organization and technology. The distinguishing feature of this method is the comparison of storage organization levels, data media, and characteristics of physical storage and stored files.

Keywords: Telemedicine System; Data Storage System; Direct Attached Storage (DAS); Network Attached Storage (NAS); Storage Area Network (SAN); Fabric Attached Storage (FAS); Network Direct Attached Storage (NDAS); Virtualization Technology; Stratified Structure.

1. Introduction

Addressing the issue of ensuring prompt and timely access to all necessary and, very importantly, reliable information about patients, which can be solved by creating a special medical profile, is one of the critical conditions for improving the quality of medical services [1]. This will obviously have a very positive impact on all services, hence improving the quality of life in society and contributing to the extension of this life. Another characteristic feature is that, in practice, the medical subject area is characterized by a geometric increase in the volume of various data used in medicine [2-4]. It should be understood that such an array of information is formed from various data, ranging from archival materials and medical history to the results of ongoing analyses, images, and examinations [5, 6]. Information about drugs and specific methods of their use is available in a huge database [7, 8]. In general, within the framework of medicine, a prodigious amount of data must be considered one way or another during ongoing manipulations, which means that this data must be stored and actively used in developed medical systems. It should also be understood that the increase in the amount of targeted information within telemedicine systems (TMS) is caused by the need to ensure security and confidentiality, which also requires the use of a large number of resources [9–11]. For example, a method can handle security issues in electronic patient health information/records and other telemedicine applications. This method behaves similar to watermark methods and provides secure and protected transmission of medical information [12].

Thus, all the above factors and reasons confirm the general relevance of the need to process large amounts of data and create all the necessary support for its storage [13–16]. However, recent studies show that the amount of medical data doubles every two years. Considerable prospects for improving the healthcare system for the better are opening up through the development of modern information technologies, among which one can note cloud storage of information and artificial intelligence, which are used to solve various typical problems. As a result, these tools create the prerequisites for organizing such complex information systems.

Data storage systems (DSS) act as a special engineering solution that ensures the formation of a developed information storage infrastructure [17, 18]. An important advantage of these systems is that they can connect various external storage devices of different physical natures [19]. At the same time, it is worth noting that, in practice, it is extremely difficult to ensure the integrated management of DSS resources [20–22]. In particular, the most important element is to consider the large number of factors presented below:

- Heterogeneous nature of storage systems. In particular, they may differ in terms of their physical nature and architecture;

- The presence of various protocols and options for requirements in the field of information storage;

- Absence of a single accepted algorithm that would determine the features of the organization of data storage with different requirements;

- The need to ensure the parallel execution of DSS processes, particularly data storage and information recovery, without stopping the execution of current tasks.

It should be noted that, to date, researchers have managed to develop a certain set of models that can be used to solve problems related to the management of DSS resources. These include shared storage models, a three-level model, a model for organizing interaction between various components, and many others [23, 24]. Furthermore, it is important to highlight that they do not provide a full-fledged possibility of accounting between different levels of information storage [25]. Thus, if we consider all the existing problems in organizing data storage, it becomes clear that it is extremely urgent to develop new models that will help successfully solve all these problems. To solve this problem, it is necessary to create a model that can account for the shortcomings of existing models. Such a model of resource management will be based on the methods of system analysis, probability theory, random processes, and mathematical statistics.

In addition, it is important to realize that society today has a skeptical attitude toward telemedicine systems. Research by Gierek et al. [26] describes a study in which more than 75% of people who had used telemedicine technology said that the technology was ineffective and that they were dissatisfied. We hope that our proposed model will increase the level of social trust in telemedicine systems. Telemedicine has been shown to extend care to populations that would not otherwise have access, including patients with cancer, who live a long distance from their oncology care, or who experience financial burden due to travel expenses and time away from work required to be seen in person. The telemedicine concept provided facilities for regular partner meetings, the possibility for discussions concerning currently treated patients, and the opportunity for patients' families to join online meetings and discuss treatment options with all partners [27, 28].

The main issues with current data storage technologies are the irrational usage of physical storage resources as well as the lack of capacity consumption forecasting capability. This paper aims to conduct comprehensive research related to the development of a multi-layered model of data storage for complex telemedicine systems to ensure the most complete use of their capacity and its timely expansion. To achieve the goal of this research work, the main features of the development of data storage technologies are systematically analyzed, and the task of forming an improved model of TMS physical storage is also solved.

2. Peculiarities of Storage Technologies for Target Data in Large Information Systems: Current State

There is no doubt that in the modern world, the development of information technologies leads to the constant formation of new opportunities for society, while specialists must solve more advanced and complex tasks related to the implementation of such information processes [29]. Information storage is one of the most critical information processes in this case since a qualitative solution to this problem ensures the most reliable operation of the information system and, hence, makes it possible to fully meet the existing needs of users. The main difficulty in organizing the storage of information is that the data may very often be needed both now and in 50 years; that is, their deletion is excluded, which necessitates the development of specific methods for storing such data.

In terms of medicine, the history of the development of procedures for the creation of drugs and the treatment of diseases, the anamnesis of patients, and a lot of related data are striking examples. The current legislation determines this differently in different countries, but the typical period of information storage is at least 25 years. Obviously, such a task requires the involvement of a huge amount of information resources. At the same time, it is essential that the existing DSSs ensure the safety of all data [30-32]. Noteworthy, the development of information technology implies analog and digital storage of information, which acts as an extremely complex and multi-layered process characterized by a multi-layered life cycle:

- Forming the need to store information at the source;

- Creating a specific message, that is, transforming existing information;
- Recording this message to a specific medium using a specific method;
- Storing this medium for the required time period;
- Organizing information reading from the medium;
- Interpreting the message, that is, providing the reverse transformation of the encoded data into the original message.

While considering the existing data storage technologies, their characteristic feature will be the wide variety of physical nature of such information media; therefore, the number of ways to access data will rapidly increase. Moreover, special DSSs are used to store colossal data arrays, acting as a set of software and hardware tools that allow recording, storing, and converting information. In most cases, the DSS consists of specific physical storage, a system that provides data backup, and special software tools that provide access to information and complex work with it. There is a trend that an increase in the total amount of information to be stored which indicates an increasing need for the creation of special information infrastructures with varying levels of complexity.

Ensuring a high level of stored information reliability requires the use of a special architecture and special technologies [33-35], the main of which are presented in Table 1. It should be noted that the results presented in Edelson [36] are outstanding.

Table 1. Technologies for organizing data storage

Technology	Essence	Features
DAS (Direct Attached Storage)	The storage device is connected directly to the server or the workstation	In essence, a DAS system is an off-server disk basket with hard drives. It is characterized by several positive and negative features. Easy to deploy and administer at a low cost, the DAS system provides high-speed exchange between the disk array and the server. However, firstly, it has low reliability: if the server to which this storage is connected fails, the data is no longer available. Secondly, it is characterized by a low degree of resource consolidation – the entire capacity is available to one or two servers, which reduces the flexibility of data distribution between servers. As a result, either more internal hard drives must be purchased or additional disk shelves must be installed for other server systems.
NAS (Network Attached Storage)	A stand-alone integrated disk system, in fact, a NAS server, with its own specialized OS and a set of useful functions for quickly starting the system and providing file access.	The NAS system is connected to a conventional computer network and provides a quick solution to the problem of lack of free disk space available to users of this network. It is characterized by a number of positive and negative features. A fairly cheap NAS system provides the following: the availability of its resources for individual servers and for any computers in the organization; ease of sharing resources; ease of deployment and administration; and versatility for clients (one server can serve MS, Novell, Mac, Unix clients). However, most low-cost NAS servers do not provide a fast, flexible method of accessing block-level data inherent in SAN systems.
SAN (Storage Area Network)	Special dedicated network connecting storage devices to application servers	The SAN system is characterized by several positive and negative features. SAN architecture provides high reliability of access to data located on external storage systems; centralized data storage; convenient centralized switch and data control; independence of the SAN topology from the used DSS and servers; transfer of intensive I/O traffic to a separate network, offloading the LAN; high speed response and low latency; scalability and flexibility of the SAN logical structure; the ability to organize standby, remote DSS, and a remote system for backup and data recovery; and the ability to build fault-tolerant cluster solutions at no additional cost. However, the SAN system is characterized by higher cost, difficulty in customization, and more stringent requirements for component compatibility and validation.
NetApp FAS (Fabric Attached Storage)	Alternative to DAS, structured and switched multi-protocol object-oriented data storage system	The FAS architecture provides a convergence of two technologies, block and file. Convergence is achieved by introducing an intermediate layer comprising objects into the file system. The FAS storage system, which has a modular structure, is characterized by several features. The DSS can work as a hybrid system or an all-flash system, depending on the needs. A high degree of integration with cloud technologies makes it possible to use external resources to protect stored data and configure equipment. The flexibility and ease of use of the hardware, together with high storage capacity, ensures network operation stability, and the scalability of the model enables the gradual increase of the amount of stored data as needed. A special System Setup utility reduces the speed of the DSS configuration by a factor of three. The NDO and QoS technologies are supported using the on Command module, making it possible to quickly integrate the DSS with various business solutions.
Unified Storage	An all-in-one storage solution that adds both NAS and SAN connectivity to NetApp FAS storage	The modular design of this storage system provides extremely wide application flexibility with the ability to gradually increase the number of storage systems in the "basic" configuration as the need arises. The NAS system, as an integral element of Unified Storage, provides cross-platform sharing of file access, and the SAN system provides distributed access to devices at the block level. Architectural features of this system provide cost savings, reducing operating costs, and increasing disk space utilization, for example, available free space remains centrally consolidated and available to applications, rather than being spread across several heterogeneous storage systems. Initially "monolithic basic" systems are much faster and more perfect than "modular" ones, but can be trivially "unaffordable", especially if all the features available in them are simply not needed at the current moment.
NDAS (Network Direct Attached Storage)	A technology enabling the connection of all digital devices (HDD, ODD, Memory, tape drives) to standard Ethernet networks. All network users and services can control and share these devices.	The NDAS technology organizes a direct connection of a physical storage medium with each client via an Ethernet network, thus eliminating performance-limiting factors (load on the CPU and network). NDAS achieves a very high level of performance due to the unique protocol and controller efficiency. NDAS uses its own protocol, but NDAS-NAS can be used by all applications running in a TCP/IP environment.

Among the main results of a comparative analysis of the data storage technologies presented in Table 1, which reflect the researchers' desire to find ways to improve them, the following should be noted first of all:

SAN is always better than NAS and DAS, but its cost is very high. In a SAN, data storage is physically or logically divided into multiple disks, which can only be accessed by a specific computer. Moreover, each computer in the network considers the storage as a DAS. Unlike NAS, SAN does not require a dedicated computer to run, which is a huge advantage that makes it possible to use higher speeds. In addition, the SAN is unaware of files; file operations are performed on servers connected to the SAN, and the SAN operates in blocks, such as a large hard drive. Ensuring that any server under any operating system can access any part of the disk capacity located in the SAN is an ideal solution for the operation of a SAN. SAN terminators are application servers and storage systems (disk arrays, tape libraries, etc.). SAN provides better performance than NAS because each logical or physical part of the hard drive is accessed from different computers. SAN offers dedicated links to various computers, which ultimately provides better speed and reliability than DAS. Due to storage technology limitations, DAS is currently being used less and less. A simple NAS in terms of configuration may be the best solution for choosing a storage system if very high speeds are not required [37].

The cutting-edge solutions that fall into the category of FAS are better than the technologies that appeared independently of each other and exist almost in parallel, such as NAS and SAN, which form the basis of FAS (primarily, this refers to unified storage systems). The advantage of FAS storage is that it provides (based on the concept of object memory storage) synergy between NAS and SAN technologies. FAS enables the consolidation of metadata into a centralized repository, which provides distributed access to multiple servers. Compared with competitors, models such as NetApp FAS have a 45% increase in capacity [38].

NDAS technology is a revolutionary solution in the field of next-generation storage systems. It is easy to use and does not require any special knowledge of networking, resource sharing, TCP/IP adjustment, or DHCP configuration. When comparing parameters such as performance, reliability, scalability, or availability, NDAS outperforms all other NAS, which have more complex architectures, rely on server structures, and have a more complex data processing scheme. NDAS is 5–6 times faster than other NASs on wireless or wired networks. Moreover, the performance of competing infrastructures, whose administration costs increase over time, significantly decreases with an increase in the number of users [39, 40].

The fundamental factors for these technologies are special redundant arrays of independent disks, which in essence act as a set of hard disks, the operation of which is additionally synchronized. This approach is used to improve DSS reliability and fault tolerance. In addition, it is worth noting that the architectures of such storages described above, which are based on different technologies, will differ in terms of the specific way the information storage is connected to the platform. It is implicit that each approach will be characterized by its advantages and disadvantages, which are presented in Table 1. The choice of each option is determined by the conditions of the specific problem to be solved.

When analyzing modern practice, it can be noted that approaches that imply the implementation of virtualization of storage systems have become extremely widespread. The essence of such a virtualization technology is reduced to organizing the logical provision of storage capacity, which means that within the framework of this approach, when providing information resources, one can abstract from the current physical implementation of the system. The use of virtualization implies that the storage structure and the DSS itself are hidden from the external environment while accessing the logical storage pool. The DSS layer manages the physical location of the information.

At the same time, it should be understood that when working with information during the operation of a virtual DSS, it will imply the transformation of the logical representation of data into physical addresses within specific drives. This implies constant work with sets of metadata, which act as a collection of additional information required by the operating system to correctly work with the available information.

Metadata plays an important role in data management. Specialized metadata allows the analysis of information about stored data blocks, files, and how they are used to form a file. Simultaneously, they describe the form of data and do not contain information about the content of the information. Organizational metadata, on the contrary, carries a semantic load. They allow for the analysis of the content of stored information, including its author, date of recording, and other additional information. The use of information virtualization makes it possible to successfully solve the problems associated with ensuring the unity of data, i.e., their consolidation, even if the data are placed on different variations of media. Moreover, this approach implies the practical implementation of information lifecycle management functions, which means the creation of multi-layered storage that combines a large number of different media. In addition, this approach helps implement automated information movement.

It should be understood that in connection with the formation of such storage networks, it became necessary to streamline various storage devices; hence, special protocols have been formed that can be used to describe specific typical solutions for the virtualization of information storage systems. Using such models, it is possible to describe various functional levels and properties of the storage system, neglecting the specific structure or features of the physical implementation of the system. Upon closer examination, it becomes clear that because of the shared storage model, a description is formed for the transformation processes of data representations that are insignificant in volume, including database files and tuples. If large data representations are considered, volumes or logical devices will already be used. The DBMS or file system enables the organization of these processes.

The most important advantage of virtualization technologies is the possibility of presenting heterogeneous media as a single information storage. In addition, from a logical viewpoint, DSS can be considered a special hierarchical structure that contains specific physical data storage and various means for providing access to it. Figure 1 shows the specific levels of DSS in more detail, regarding the specifics of the storage process life cycle.
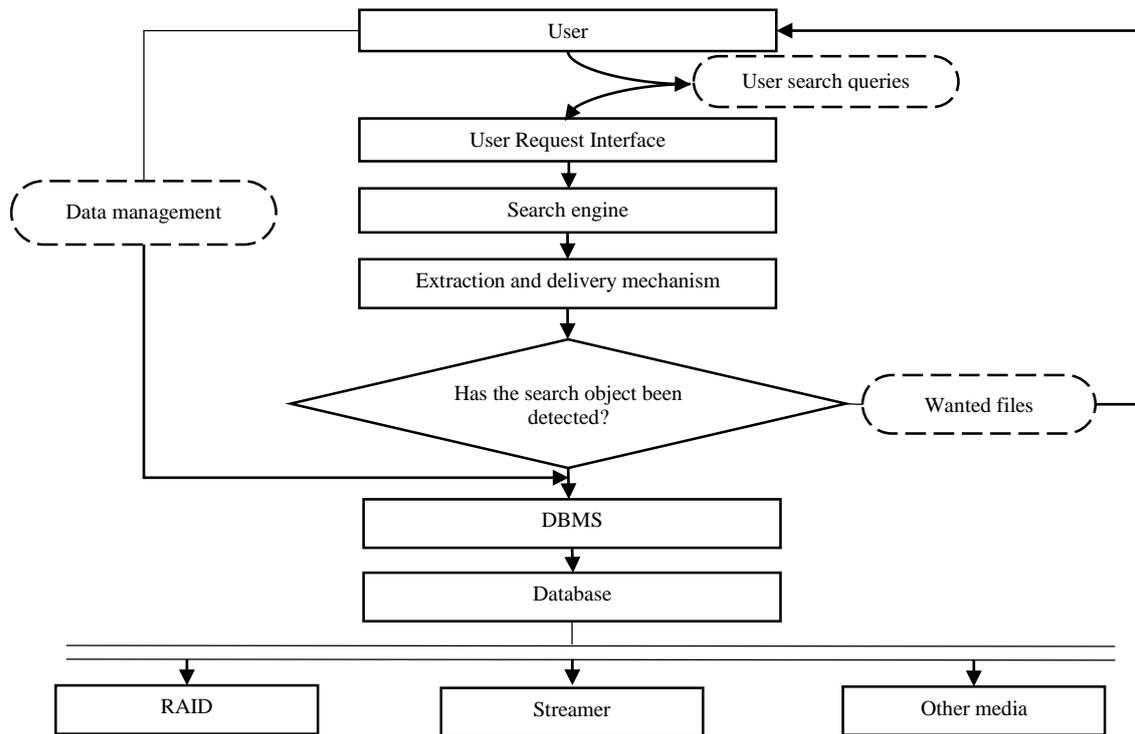
Figure 1. Hierarchical DSS structure with specific data storage

Noteworthy, most of the studies conducted focus on the process of data management. Studies in the field of the physical nature of data storage constantly explore the possibility of practical implementation of promising technologies for recording information, such as tungsten disks and biological recording. Furthermore, an increasing number of advanced approaches to data recording are constantly being developed. The variety of modern media is completely determined by the physical nature of storage, the technologies used, and the material of the particular media. Undoubtedly, all these factors determine the service life, and it is essential to select them depending on the type of information they store, since various information variations will imply a different storage period for this data. It is worth noting that nowadays it has been possible to achieve tremendous progress in the matter of a significant increase in the density of information recording. This has dramatically expanded the potential range of services that can be provided to ordinary users and information system designers. Simultaneously, if we return to the examination of Figure 1, it will be possible to find that the existing lower-level problems are given much less attention.

Most of the ongoing studies aimed at improving existing technologies for storing information state the need to further reduce the cost of storing a unit of information, which is a bit of a main goal. It is quite clear that price reductions can be achieved using organizational and specific technical approaches. The organizational ones include the division of the stored content depending on the specific type of information, after which different versions of the information will be stored on different media. As a result, the use of this approach can significantly save overall storage space. Moreover, the data storage process itself implies the use of a certain encapsulation chain: the minimum storage unit (MSU) → a physical data block (PB) → a logical data block (LB) → a file. The hierarchical structure of this chain and the relative comparative sizes of its blocks are illustrated in Figure 2.
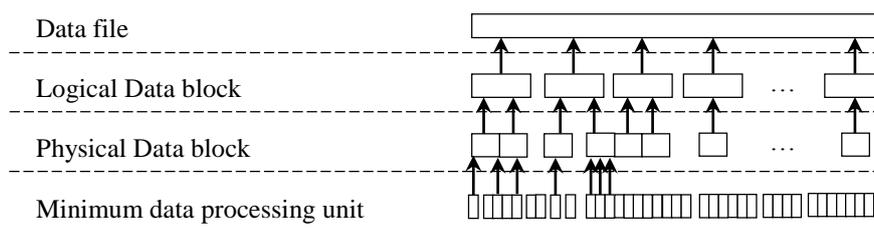
Figure 2. Illustration of the hierarchical structure of the data encapsulation chain in the implementation of data storage and the relative comparative sizes of its blocks

It is proposed to dwell on the essence of this chain in more detail. It consists of the following: Here, the minimum storage unit is the minimum physical area of the storage medium, the main property of which is the ability to be in one of several stable states, the number of which determines the number of data bits stored: if the number of states is 2, the MSU can store 1 bit of data; if this number is 4, 2 bits can be stored; and in the case of 8 states, 3 bits can be stored, etc.

Further, at the next level, the minimum amounts of data are combined into special physical blocks. A special structure is considered a physical data block, which helps combine different minimum units of information storage. The specific address of the information location is the most important characteristic of such a structure. The size of the physical block can be adjusted depending on the type of information being recorded and the requirements of the storage system. For example, it can also have a traditional user data format of 512 bytes or 4096 bytes for the Advanced Format. In turn, the logical block, the size of which is set during formatting and depends on the capabilities of the file system used, will be a special structure that describes specific addresses of physical data blocks. The file system manages files, which are a logical structure of pooled data with a semantic load and contain all the addresses of logical blocks. Modern file systems require the ability to simultaneously work with several files up to 16 GB in size. In the process of forming a file using addressing, the data bits are combined into a file (see Figure 3).
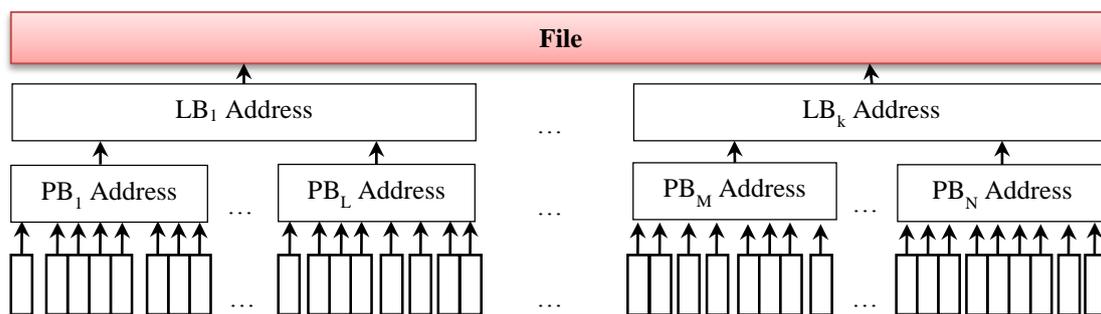

Figure 3. A file structures

It is also worth considering that the implementation of data storage procedures requires the availability of special intellectual, informational, and physical resources. In turn, the process of providing access to information will imply the existence of an appropriate logical structure that will help search for and work with data. Relational and non-relational databases act as a similar structure within modern data storage systems. These databases are inherently a special logical data structure, enabling the establishment of all necessary links between information items.

Relational databases have become widespread because of their rigid set structure, which makes it easier to work with information; however, relational databases lose their effectiveness regarding working with large amounts of data. In this case, the use of non-relational systems begins, providing the ability to work with clusters, which allows for the increase of hardware storage resources without any problems in the absence of a single scheme. This means that new fields can be easily created in the database without changing its structure. Thus, this approach is characterized by much greater flexibility.

To ensure interaction between the user and the database, a special information retrieval system is used. Based on the results of the analysis of existing modern information technologies used for data storage, it was possible to determine that the operation principle of the ANSI/SPARK model and various virtualization technologies act as an approach that allows physical data storage to be considered as a separate structure that does not directly impact the presentation of data. Furthermore, the information management process, like itself, implies the implementation of direct work with metadata. It also becomes obvious that physical storage can contain a huge number of different media, which differ significantly in their characteristics. In this case, the information distribution process is completely determined by the specific characteristics of the information. The need to reduce the cost of a bit of information is becoming increasingly obvious due to the ever-increasing amount of stored information. It is important to understand that cost reduction can be achieved not only by increasing the recording density or guaranteed storage time but also by more efficiently organizing the distribution of available physical resources, particularly capacity.

The main goals of the implemented physical storage management systems are to maximize the exploitation of the available storage capacity and to realize the potential for increasing this capacity. Thus, it is possible to identify a specific list of tasks that ensure the development of the management system:

• Developing new, more efficient ways of distributing information within data storages for more optimal use of capacity;

• Forming accurate forecasts of capacity consumption, which will enable a rapid increase in this parameter.

Along with this, it becomes clear that it is necessary to create increasingly advanced file systems capable of working with huge blocks of data while ensuring extremely high density for the recorded information. In turn, certain requirements are put forward in relation to control mechanisms, the main task of which is to promptly provide all the

necessary information to the user. Thus, undoubtedly, at the current stage, new and more promising solutions are required in the field of organizing the storage of a large amount of targeted information; hence, completely new models are required to implement such solutions. Stratification is one of the promising solutions, which implies the creation of a family of models, each of which will determine the behavior of the system at different levels, which will solve the problem of finding a compromise between the simplicity of description and the need to consider the diverse behavioral characteristics of complex systems. As part of the stratification for each level, it will be possible to define some specific features and variables to describe the operation of the system.

3. Multi-Layered Model of the Physical Data Storage Complex

When considering almost any modeling methodology, it is possible to find that they contain a special system architecture, which helps determine specific methods of analysis and system design; moreover, an accurate description is given here for a set of strategies for using such systems. The main characteristics of the architecture of such systems are specifically defined layers, formal interfaces between each layer, and hidden elements and components. Thus, based on such features, it can be concluded that this concept implies a preliminary division into layers and the use of appropriate methodologies and technologies for structural modeling.

It should be noted that the practical introduction of modern virtualization technologies provides extremely wide opportunities for implementing various architectural solutions, and the main advantage is the absence of the need to be tied to a specific type of equipment. Such advantages make it possible to consider information storage as a separate structure with certain properties and characteristics. Furthermore, it should be emphasized that such models (first of all, this applies to such stratified models as: SNIA Shared Storage Model, the three-level ANSI for DBMS model, the model for interactive components (MIC), and the Open Systems Interconnection Reference Model (OSI-RM) have a single concept, starting with operating with binary data, implementing physical and logical addressing, establishing communication between end points, and ending with data presentation and application operation. A comparison of the information areas of these models is given in Table 2.

Table 2. Comparison of the levels of the data storage models

No.	SNIA Shared Storage Model	Three-level ANSI for DBMS	MIC	OSI-RM
7	Applications	User level	*Conceptual data* (does not correspond to the applied layer of OSI-RM)	Applied
6			Query language	Presentation layer
5	File virtualization	Conceptual level		Session
4			Data access	Transporting
3	Block virtualization		File systems	Networking
2		Physical level	Memory	Canal
1	Block sub-system		High-capacity buses and external drives	Physical

It should be noted that the existing models have no detailed description of the recording and physical storage of data bits, which is necessary for describing the storage process. A distinctive feature of the proposed multi-layered data storage model is the comparison of storage organization layers, data media, and the characteristics of physical storage and stored files. A model is proposed that considers both the characteristics of the data and the organizational, technical, and technological features of the data storage process. The model has a stratified structure that combines storage layers, types of media (i.e., technologies that implement this process), and data characteristics.

Consider a function that enables the formal representation of any hierarchical data storage system:

$$D = \varphi(C, R, F) \tag{1}$$

where D is a multi-layered data storage system; C – the data storage structure, defined by the layers at which the stages of the life cycle of the data storage process are implemented: $C = \{m, n\}$, where m is the number of storage layers; n_i – the number of volumes (media) on the i–th storage layer;

R – a set of temporary, capacitive, energy resources required for the implementation of the storage process;

F – incoming data stream for recording with subsequent storage, $F = \{t, S, f, \lambda\}$, where t is the time of data storage in years; S – the size of the logical data block in bytes; f – the file size in bytes; λ - data access frequency, number of requests/hours.

In this case, the immediate task of the study will be to create special models for managing DSS resources, which will be characterized by the following opportunities:

- Accounting for the main features of the specifics of the implemented information process related to information storage;

- Accounting for the features of a particular file structure;

- Managing the use of the available physical storage resources;

- Creating specific forecasts for the consumption of storage capacity, which makes it possible to quickly increase such volumes in practice. Such an approach will significantly increase the reliability of the information storage process within the DSS.

Thus, it is proposed to use a multi-layered model, the characteristic features of which will be disclosed in more detail later. This model enables the description of the existing information processes within the digital environment; a stratified structure can be noted among the features of the models.

Consider the following specific prerequisites for the formation of stratified models:

- Information processes are implemented using special computer systems that are managed as a single whole. Moreover, the systems themselves consist of a combination of hardware and software.

- The formation of the information process description is based on a specific sequence of functional tasks to be solved, whereas the specific method of their implementation is unimportant.

- To represent the existing functional tasks, hierarchical layers are used, united by special links;

- Information processes are supported with the help of service information.

Note that stratified models are formed on the basis of the principle of dividing the information process into a sequential chain of interconnected functions, while each of these functions will be at different layers of the hierarchy. It is important to understand that such levels will have a vertical decomposition, which allows us to say that they are a stratified description.

One of the characteristic features of the multi-layered data storage model under consideration is the possibility of practical comparison of different layers of storage organization and data media. In addition, the specific characteristics of physical storage and saved files are considered. The use of this model will allow for consideration of all organizational, technological, and technical features of the information storage process. As noted earlier in this research, the model itself will have a stratified structure that integrates multiple storage layers and media variations. At the same time, within the framework of the model under consideration, it is proposed to use only physical data storage, and the architecture should consider the combination of the following features:

- In its nature, the information storage process can be considered as a process of transferring information in time, after which the essence of this information is conveyed to the specific user who requested it after a specified storage time. Thus, ensuring a stable and safe state of information is the most important parameter.

- Storage functions should be considered simultaneously at several layers. Moreover, the files on the media act as logical blocks of information, and the file system sets their specific characteristics. In turn, logical blocks are structured as physical blocks, which are the minimum units of storage. The process of organizing information storage involves the placement of this information in different storage layers. It is also required to provide comprehensive control of information migration depending on the specific layers of physical storage.

- Changing the state of the MSU is the primary step in data recording.

- To implement information storage in physical storage, special physical resources are required. The saving of these resources is ensured by the competent administration of this information.

- If it is necessary to use systems for the long-term storage of information, it may be necessary to create specific file systems within which the features of such data will be considered.

It is proposed to perform the operation of vertical, horizontal, and dynamic data placement mechanisms sequentially. The first stage is data allocation to storage levels depending on the metadata containing the type of files to be stored. Each storage level assumes a certain data storage time and frequency of access. The lower the level, the longer the storage time and the lower the frequency of data access. The type of files to be stored is purely subjective and is determined by the information owner or technician.

In the second step, once the storage tier is determined, the horizontal distribution of files to the selected tier is performed based on the size of the data file. At the third stage, the data are migrated to the storage levels depending on the frequency of access. The frequency of access is determined on the basis of statistics for a certain period. Threshold values can be taken into account when overcoming which the data migrate to a more suitable storage level. Obviously,

the migration mechanism allows overcoming the disadvantages of subjective choice of the type of files to be saved when implementing the first stage (distribution of data to storage levels when writing). The aggregate of the above mechanisms allows management of the storage capacity and rational utilization of the media. It should be noted that the above mechanisms are based on the analysis of the metadata of the stored files.

The architecture of the described model is presented in more detail in Figure 4. The presented tiered storage model differs from other open systems models by a stratified description that combines tiered storage organization, data storage technologies, and characteristics of physical storage and stored files. This allows storage designers to obtain solutions with specified properties across a variety of storage options.
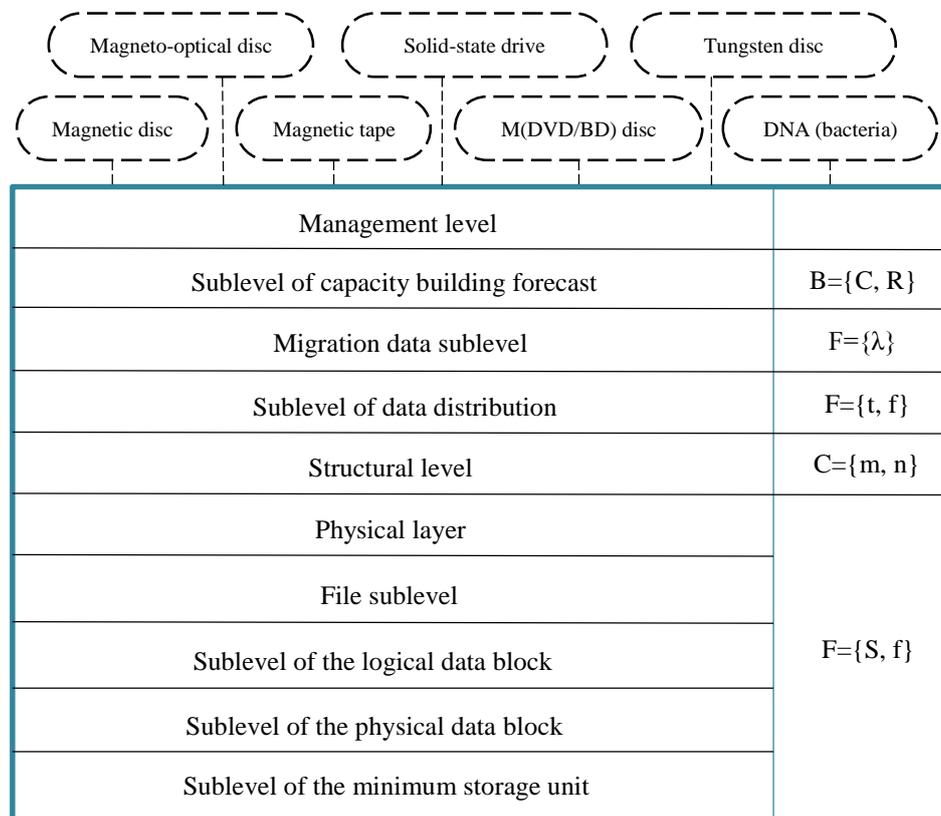
Management level	
Sublevel of capacity building forecast	B={C, R}
Migration data sublevel	F={λ}
Sublevel of data distribution	F={t, f}
Structural level	C={m, n}
Physical layer	
File sublevel	
Sublevel of the logical data block	F={S, f}
Sublevel of the physical data block	
Sublevel of the minimum storage unit	

Media types (top of figure): Magneto-optical disc, Solid-state drive, Tungsten disc, Magnetic disc, Magnetic tape, M(DVD/BD) disc, DNA (bacteria)

Figure 4. Architecture of the proposed multi-layered data storage model

As already noted in this research, the strata of a multi-layered data storage model act as levels characterized by vertical decomposition and describe a specific sequence of data storage organization. The functions of the sublevels are described in more detail below. There are three levels of data storage in the proposed model. The first of them – physical, is shown in Table 3.

Table 3. Functions of the physical layer of the multi-layered data storage model

Physical layer of the multi-layered data storage model		
Sublevels	**Functions**	**Note**
File sublevel	Defining specific addresses for bits and physical and logical blocks	In fact, at this level, all available data bits are logically united into a file.
Sublevel of the logical data block	Forming a logical data block	Physical data blocks undergo encapsulation and transformation into logical data blocks.
Sublevel of the physical data block	All minimum storage units are united, after which the data physical blocks are formed	The main result of this layer is MSU transformation into physical data blocks by encapsulation.
Sublevel of the minimum storage unit (MSU)	Changing the minimum storage unit state	Within the MSU sublevel, functions aimed at providing support for the stable storage of data units are implemented. All minimal units of data are capable of being in a steady state for a certain time.

Furthermore, Figure 5 describes the essence of the layered structure of data storage in more detail, showing the structural level of the model. Data storage will have a matrix structure in mathematical form, where each cell of the matrix is a volume of the corresponding storage tier designed to store data with certain metric characteristics.
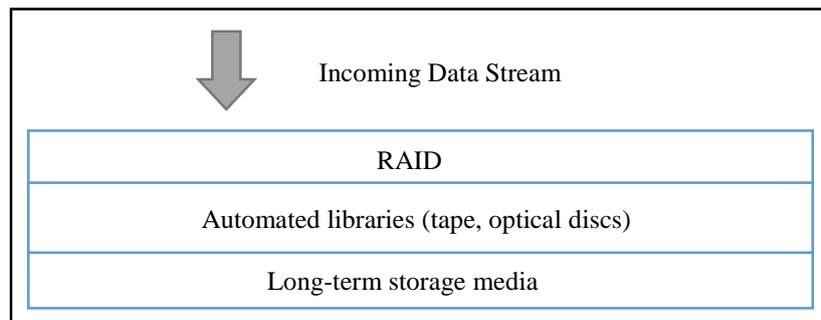
Figure 5. Layered structure of the data storage

It is worth considering that each level of information storage will imply the use of specific storage technologies. Moreover, before the next recording of information within a particular level, data are analyzed to select the optimal file system for processing. The flowchart of the workflow is shown in Figure 6. If the management layer is considered in more detail, it contains a detailed description of the management processes on the capacity of physical data storage. A description of the specific functions of the three main sub-levels is given in Table 4.

Table 4. Management-level functions

| Management level of the multi-layered data storage model		
Sublevels	**Functions**	**Note**
Sublevel of the capacity building forecast	Forming specific forecasts to increase storage matrix cells	The general growth in the amount of data will lead to the need to increase the storage capacity within any system. It should also be considered that, in parallel, the intensity of the incoming data fairly significantly changes in the DSS. This necessitates the creation of forecasts to increase capacity. When forecasting, factors such as DSS use intensity, functional load, type of stored data, and maximum capacity of each specific cell are considered.
Data migration sublevel	Decision-making regarding the need for data migration	Data migration by levels is carried out depending on the frequency of access to specific information; a dynamic allocation mechanism is used for migration.
Data distribution sublevel	Distribution of data among cells of the storage matrix	The characteristics of the incoming data stream determine the features of the process of distributing information to different levels. Various mechanisms are used here, among which the mechanisms of vertical and horizontal placement can be distinguished; each mechanism is characterized by its own features.
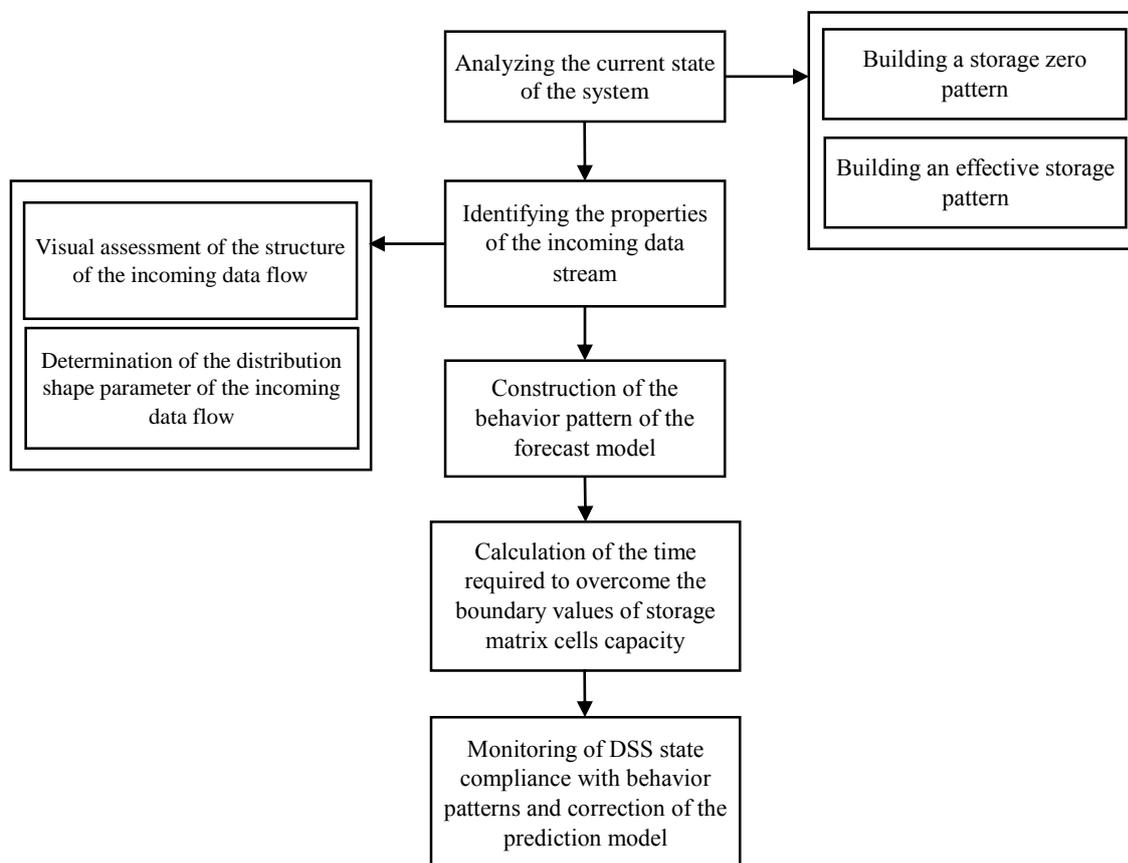
Figure 6. Flowchart of the methodology for forecasting capacity expansion.

4. Results and Discussion

Based on the research results, the proposed multi-level data storage model with a stratified structure is discussed. This model makes it possible to successfully combine different levels of storage organization, media variations, and data characteristics. The process of this model functioning is described in more detail in Figure 7. In short, the essence is that at the primary stages, before recording to a physical medium, the file characteristics are analyzed, after which a specific decision is made to record the file to a certain level. After recording, data are constantly collected on the frequency of access to each file, which enables them to be migrated based on this indicator.
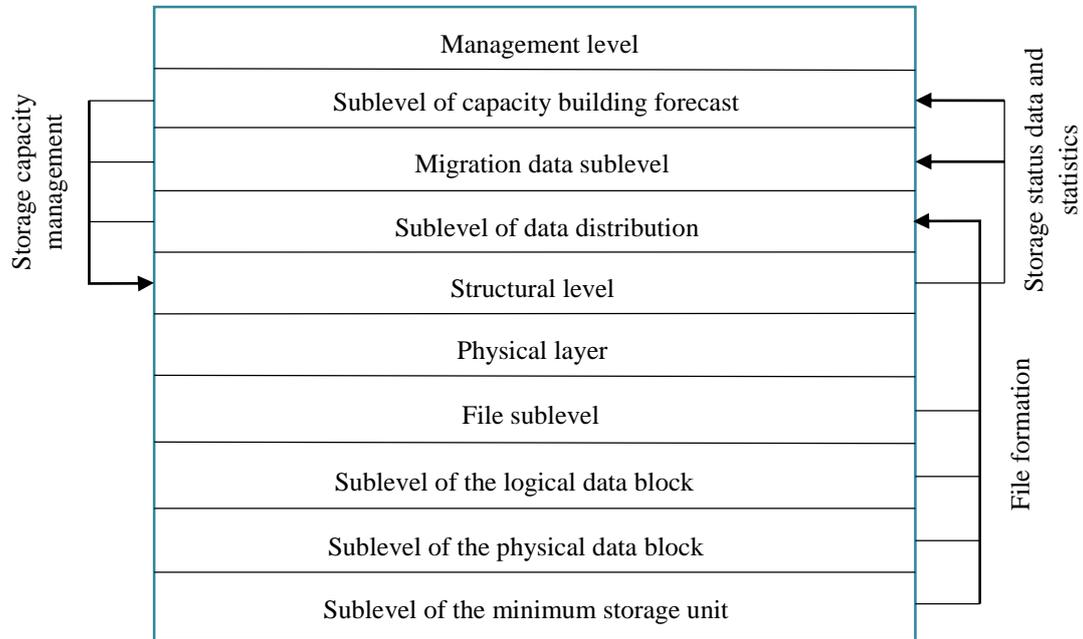

Figure 7. Storage process in the framework of the proposed model

A demonstration experiment using, for example, such a tool for solving the problem of placing objects in accordance with the required parameters, as the apparatus of the Kohonen neural network, showed that the model of resource management of data storage systems proposed in this research is operable (which is illustrated by one of the fragments of the obtained test results in Figure 8) and makes it possible to solve the problem of rational use of the DSS physical resources.
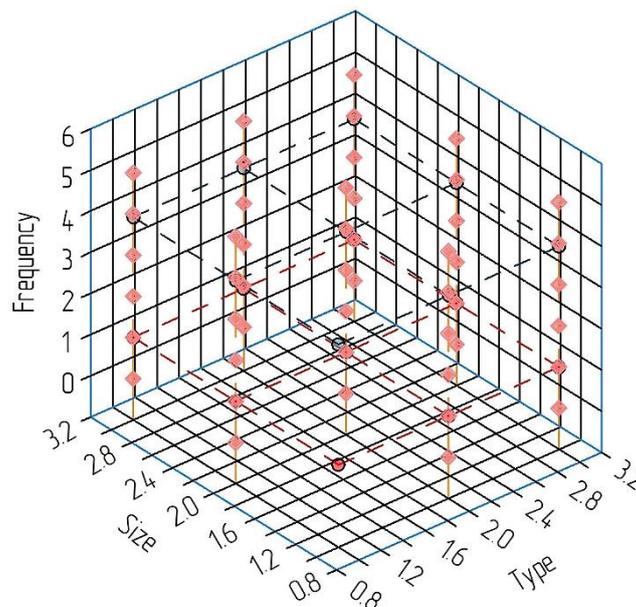

Figure 8. An example of visualization of boundary values of file access frequency

The demonstration of using the Kohonen neural network method to distribute a set of files into storage matrix cells after training the network and normalizing the initial metadata. The experiment considered an incoming stream of files

with different characteristics, such as estimated storage time, size, and frequency of data access. The stream was generated on the basis of the analysis of files taken from the experimental file server. The composition of these files was assessed, and a sample was taken for the experiment, the composition of which corresponds in percentage to the composition of the files on the considered arbitrary file server.

In this study, the performance of various models already in existence has been examined, such as SNIA, ANSI, MIC, and OSI-RM. These models have a unified concept, starting with working with binary data, realization of physical and logical addressing, establishment of communication between endpoints, and finishing with data representation and application operation.

The peculiarity of the proposed multilevel data storage model is the mapping of storage organization levels, data carriers, and the characteristics of physical storage and stored files. The proposed model was developed considering the characteristics of the data and the organizational, technical, and technological features of the data storage process. The proposed model will unify the levels of organization of the storage process, reduce the number of physical resources, and improve the quality of telemedicine systems.

The main difficulty in implementing this approach when solving the problem of placing objects in accordance with the required parameters is the choice of parameter classification and normalization. The dynamic distribution of files by levels of the storage matrix consists of analyzing data on the file access frequency and making a decision on the need for file migration. The boundary values for the access frequency, at which files are migrated by levels, must be defined by the storage system administrator.

5. Conclusion

The digital healthcare system implies the use of cutting-edge information technologies to support an increasingly efficient and cost-effective medical practice. There is a trend toward similar technological transformations worldwide. Its main tasks include a general increase in the availability, quality, and comfort of medical care for the population and the formation of the most accurate and in-depth results of medical diagnostics. This paper defines a specific research task that is reduced to the implementation of specific DSS resource management models aimed at providing a solution to the problem of forecasting the consumption of these resources. The DSS management system, whose main purpose is to ensure the fullest use of storage capacity and its timely expansion, was considered in detail. The potential benefits of applying the model presented in this paper are its application to federated storage resource management, given the relationship between storage tiers, characteristics of stored data, and write technologies. A multilevel model of data storage is proposed that combines the levels of organization of the storage process, the technologies that implement this process, and the corresponding data characteristics. The functions of the levels and sub-levels of the proposed version of the multilevel data storage model and the operation of the multilevel model are described. The problems of realizing the data storage process at each level of the proposed model, which consist of saving physical resources and timely increasing the storage capacity, are considered. This development will improve the quality and credibility of complex telemedicine systems.

6. Declarations

6.1. Author Contributions

Conceptualization, A.T. and D.P.; methodology, D.P.; software, A.L.; validation, A.T.; formal analysis, D.P.; investigation, A.L.; resources, D.P.; data curation, A.T.; writing—original draft preparation, A.L.; writing—review and editing, A.T.; visualization, D.P.; supervision, A.T.; project administration, A.T.; funding acquisition, A.T. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. References

[1] Kruk, M. E., Gage, A. D., Arsenault, C., Jordan, K., Leslie, H. H., Roder-DeWan, S., Adeyi, O., Barker, P., Daelmans, B., Doubova, S. V., English, M., García-Elorrio, E., Guanais, F., Gureje, O., Hirschhorn, L. R., Jiang, L., Kelley, E., Lemango, E. T., Liljestrand, J., . . . Pate, M. (2018). High-quality health systems in the Sustainable Development Goals era: time for a revolution. The Lancet Global Health, 6(11), e1196–e1252. doi:10.1016/s2214-109x(18)30386-3.

[2] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. Journal of Big Data, 6(1), 54. doi:10.1186/s40537-019-0217-0.

[3] Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. New England Journal of Medicine, 375(13), 1216–1219. doi:10.1056/nejmp1606181.

[4] Chen, M., Hao, Y., Hwang, K., Wang, L., & Wang, L. (2017). Disease prediction by machine learning over big data from healthcare communities. IEEE Access, 5, 8869–8879. doi:10.1109/access.2017.2694446.

[5] Lebedev, G. S., Linskaya, E. Y., Tatarkanov, A. A., & Lampezhev, A. K. (2023). Recent solutions in the field of automated monitoring and quality control of telemedical services. International Journal of Engineering Trends and Technology, 71(1), 62–78. doi:10.14445/22315381/ijett-v71i1p207.

[6] Kuklin, V. Z., Alexandrov, I. A., Umyskov, A. A., & Lampezhev, A. K. (2022). Analysis of the prospects for developing storage and processing complexes for multiformat media data. Journal of Computer Science, 18(12), 1159–1169. doi:10.3844/jcssp.2022.1159.1169.

[7] Chen, X., Yan, C. C., Zhang, X., Zhang, X., Dai, F., Yin, J., & Zhang, Y. (2015). Drug–target interaction prediction: databases, web servers and computational models. Briefings in Bioinformatics, 17(4), 696–712. doi:10.1093/bib/bbv066.

[8] Li, J., Zheng, S., Chen, B., Butte, A. J., Swamidass, S. J., & Lu, Z. (2015). A survey of current trends in computational drug repositioning. Briefings in Bioinformatics, 17(1), 2–12. doi:10.1093/bib/bbv020.

[9] Hale, T. M., & Kvedar, J. C. (2014). Privacy and security concerns in telehealth. AMA Journal of Ethics, 16(12), 981–985. doi:10.1001/virtualmentor.2014.16.12.jdsc1-1412.

[10] Ahmad, R. W., Salah, K., Jayaraman, R., Yaqoob, I., Ellahham, S., & Omar, M. (2021). The role of blockchain technology in telehealth and telemedicine. International Journal of Medical Informatics, 148, 104399. doi:10.1016/j.ijmedinf.2021.104399.

[11] Tatarkanov, A., Lampezhev, A., Polezhaev, D., & Tekeev, R. (2022). Suboptimal biomedical diagnostics in the presence of random perturbations in the data. International Journal of Engineering Trends and Technology, 70(11), 129–137. doi:10.14445/22315381/ijett-v70i11p213.

[12] Rashid, A., Salamat, N., & Prasath, V. (2018). An algorithm for data hiding in radiographic images and ePHI/R application. Technologies, 6(1), 7. doi:10.3390/technologies6010007.

[13] Tatarkanov, A. A., Umyskov, L., Tekeev, R. K., & Kuklin, V. Z. (2022). Model development of universal hardware and software module for medical information system. International Journal of Emerging Technology and Advanced Engineering, 12(10), 136–146. doi:10.46338/ijetae1022_15.

[14] Xia, Q., Sifah, E. B., Asamoah, K. O., Gao, J., Du, X., & Guizani, M. (2017). MeDShare: trust-less medical data sharing among cloud service providers via blockchain. IEEE Access, 5, 14757–14767. doi:10.1109/access.2017.2730843.

[15] Alyass, A., Turcotte, M., & Meyre, D. (2015). From big data analysis to personalized medicine for all: challenges and opportunities. BMC Medical Genomics, 8(1), 1-12. doi:10.1186/s12920-015-0108-y.

[16] Dimitrov, D. V. (2016). Medical Internet of Things and Big Data in Healthcare. Healthcare Informatics Research, 22(3), 156-163. doi:10.4258/hir.2016.22.3.156.

[17] Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of "Big Data" on cloud computing: Review and open research issues. Information Systems, 47, 98–115. doi:10.1016/j.is.2014.07.006.

[18] Warren, J., & Marz, N. (2015). Big Data: Principles and Best Practices of Scalable Realtime Data Systems. Manning Publications: Shelter Island, New York, USA.

[19] Haghi, M., Thurow, K., & Stoll, R. (2017). Wearable devices in medical internet of things: scientific research and commercially available devices. Healthcare Informatics Research, 23(1), 4. doi:10.4258/hir.2017.23.1.4.

[20] Jennings, B., & Stadler, R. (2014). Resource management in clouds: survey and research challenges. Journal of Network and Systems Management, 23(3), 567–619. doi:10.1007/s10922-014-9307-7.

[21] Hong, C. H., & Varghese, B. (2019). Resource Management in Fog/Edge Computing. ACM Computing Surveys, 52(5), 1–37. doi:10.1145/3326066.

[22] Musaddiq, A., Zikria, Y. B., Hahm, O., Yu, H., Bashir, A. K., & Kim, S. W. (2018). A survey on resource management in IoT operating systems. IEEE Access, 6, 8459–8482. doi:10.1109/access.2018.2808324.

[23] Yoshida, H. (2018). Storage resource management. Encyclopedia of Database Systems, 3759–3760. doi:10.1007/978-1-4614-8265-9_1342.

[24] Grozev, N., & Buyya, R. (2013). Performance modelling and simulation of three-tier applications in cloud and multi-cloud environments. The Computer Journal, 58(1), 1–22. doi:10.1093/comjnl/bxt107.

[25] Fu, J. S., Liu, Y., Chao, H. C., Bhargava, B. K., & Zhang, Z. J. (2018). Secure Data Storage and Searching for Industrial IoT by Integrating Fog Computing and Cloud Computing. IEEE Transactions on Industrial Informatics, 14(10), 4519–4528. doi:10.1109/tii.2018.2793350.

[26] Gierek, M., Kitala, D., Łabuś, W., Glik, J., Szyluk, K., Pietrauszka, K., Bergler-Czop, B., & Niemiec, P. (2023). The impact of telemedicine on patients with Hidradenitis Suppurativa in the COVID-19 Era. Healthcare, 11(10), 1453. doi:10.3390/healthcare11101453.

[27] Calton, B. A., Nouri, S., Davila, C., Kotwal, A., Zapata, C., & Bischoff, K. E. (2023). Strategies to make telemedicine a friend, not a foe, in the provision of accessible and equitable cancer care. Cancers, 15(21), 5121. doi:10.3390/cancers15215121.

[28] Troschke, T., Wieczorek, A., Kulinski, K., Ociepa, T., Zielezinska, K., Lode, H. N., & Urasinski, T. (2023). Pediatric hematology and oncology center integrated by telemedicine: experience, challenges and first results of a cross border network. Healthcare, 11(10), 1431. doi:10.3390/healthcare11101431.

[29] Kumar, K. (2020). From post-industrial to post-modern society. In The Information Society Reader, Routledge, pp. 103-120.

[30] Akhtar, D. N., Kerim, D. B., Perwej, D. Y., Tiwari, D. A., & Praveen, D. S. (2021). A comprehensive overview of privacy and data security for cloud storage. International Journal of Scientific Research in Science, Engineering and Technology, 113–152. doi:10.32628/ijsrset21852.

[31] Chen, Y., Ding, S., Xu, Z., Zheng, H., & Yang, S. (2018). Blockchain-based medical records secure storage and medical service framework. Journal of Medical Systems, 43(1). doi:10.1007/s10916-018-1121-4.

[32] Yang, P., Xiong, N., & Ren, J. (2020). Data security and privacy protection for cloud storage: a survey. IEEE Access, 8, 131723–131740. doi:10.1109/access.2020.3009876.

[33] Cha, B., Park, S., Kim, J., Pan, S., & Shin, J. (2018). International network performance and security testing based on distributed abyss storage cluster and draft of data lake framework. Security and Communication Networks, 2018, 1–14. doi:10.1155/2018/1746809.

[34] Malav, V., & Sharma, D. A. (2018). Effect and benefits of deploying Hadoop in private cloud. National Journal of Multidisciplinary Research and Development, 3, 1057-1062.

[35] Park, J. K., & Kim, J. (2018). Big data storage configuration and performance evaluation utilizing NDAS storage systems. AKCE International Journal of Graphs and Combinatorics, 15(2), 197–201. doi:10.1016/j.akcej.2017.09.003.

[36] Edelson, E. (2004). Security in Network Attached Storage (NAS) for Workgroups. Network Security, 2004(4), 8–12. doi:10.1016/s1353-4858(04)00065-0.

[37] Salim, N. B., Zambri, N. A., Suhaimi, M. B., & Sim, S. Y. (2023). Automatic generation control system: the impact of battery energy storage in multi area network. International Journal of Integrated Engineering, 15, 208–216. doi:10.30880/ijie.2023.15.03.022.

[38] Liu, M. (2023). Fabric-centric computing. In Proceedings of the 19th Workshop on Hot Topics in Operating Systems (pp. 118–126), Association for Computing Machinery, Providence, RI, USA. doi:10.1145/3593856.3595907.

[39] Okafor, I., Ramanathan, A. K., Challapalle, N. R., Li, Z., & Narayanan, V. (2023). Fusing in-storage and near-storage acceleration of convolutional neural networks. ACM Journal on Emerging Technologies in Computing Systems, 20(1), 1–22. doi:10.1145/3597496.

[40] Zet, C., Dumitriu, G., Fosalau, C., & Sarbu, G. C. (2023). Automated calibration and DCC generation system with storage in private permissioned Blockchain network. Acta IMEKO, 12(1), 1–7. doi:10.21014/actaimeko.v12i1.1414.