DOI: 10.32347/2412-9933.2020.44.166-174
UDC 004.042

Chernyshev Denys
DSc (Eng.), Professor, First vice-rector, *orcid.org/0000-0002-1946-9242*
Kyiv National University of Construction and Architecture, Kyiv
Tsiutsiura Svitlana
DSc (Eng.), Professor, Head of Department of Information Technology, *orcid.org/0000-0002-4270-7405*
Kyiv National University of Construction and Architecture, Kyiv
Lyashchenko Tamara
Senior Lecturer of the Department of Information Technology, *orcid.org/0000-0001-9092-0297*
Kyiv National University of Construction and Architecture, Kyiv
Luzina Yuliia
Student of the IV class of the Department of Information Technologies, *orcid.org/0000-0002-8426-9431*
Kyiv National University of Construction and Architecture, Kyiv
Borodynia Vitalii
Postgraduate student of the Department of Information Technologies of Design and Applied Mathematics, *orcid.org/0000-0001-9513-92645*
Kyiv National University of Construction and Architecture, Kyiv

THE USE OF DATA WAREHOUSES ON THE EXAMPLE OF THE APPLICATION PROCESSING SYSTEM IN GOVERNMENT AGENCIES

Abstract. *This article highlights the issue of the importance of information, the need to ensure its proper storage and use. Nowadays, the question of the importance of studying the functioning of data storage facilities in one of the most common design models – Data Flow Diagram, and data storage (DS or DW) in the whole. A study of abstract devices of this information security model was carried out: types, features and types of data storage, which are used in information systems (relational, multidimensional, and hybrid repositories), from the point of view of using models for presenting data; appearance of warehouses and building rules; Letter identifiers of storages "D", "C", "M" and "T" with the help of which the type of storages is determined; features of the numerical part of the identifiers for decompositions of the first and second tier processes; mechanisms that support data retention for their intermediate processing in information systems. Transition of properties and characteristics from physical to logical representation, rationalization of data warehouses by considering the features of the logical model. In the course of the work, the peculiarities of the construction of the DFD and the reflection of the interrelationships in all the component diagrams, defined in the general rules that are valid for the two, were considered. The highlighted issues concern the elements of the diagrams, which can be freely exploited within the borders of Ukraine, as well as the elemental content of the different types of diagrams in the diagrams. The unique functionality of domestic DFD diagrams allows to use a sufficient, albeit limited, elemental structure for construction. Depending on the actual possibilities, the diagrams of the Ukrainian manufacturer are somewhat simplified and imperfect according to modern technologies, but these factors do not diminish the importance and necessity of adequate data protection at the highest level. Therefore, in this work, the main most informative moments of the use of this or that type of storage, the demand for DS, in the field, are highlighted, possible advantages and disadvantages of using physical data storage facilities and features of virtual operation.*

Keywords: data flow diagram; data storage; database; saving information

Introduction

In our time, information is the most important element of communication between people, an integral part of development and the relationship between the people of the entire planet. From the moment of awareness of the importance of information as a product, there is a question about its preservation and the appropriate use of data in later systems.

Any design of the information system (IS) – from the observation of the enterprise to the creation of a data model for implementation in a specific database can not exist without it. In this paper, the authors propose to consider an abstract procedure for storing information – data storage.

In order to get acquainted in more detail with the concept and possibilities of the data warehouse, the preservation of information on the example of the functionality of the data flow diagrams – DFD is considered.

The main purpose of DFD is to show how each job converts its input data to the output, and also to identify the relationship between these jobs. Any DFD-diagram contains jobs, external entities, data (data streams) and data storage.

Data Flow Diagram is a model of designing, graphically representing "data flows" in an information system [1; 2]. Data flow diagrams represent a modulated system as a network of interconnected beats and are used to describe document flow and information processing. Diagrams include work, external references, as well as data storage. To deal with such large amounts of information, it is necessary to use a DS.

Therefore, on the example of DFD work, the following will be considered:

1. Principles of work of the data center.
2. Features of data warehouses.
3. The field of data storage in information systems.
4. Relationships of data warehouses with other components of diagrams.

The purpose of the article

The purpose is to look at the peculiarities of saving and further use of data in the field of graphical representation of systems; connection of the information preservation process with other components in different systems.

Basic types of DS

In a broad sense, a data warehouse is a subject-oriented, integrated, unchanging set of data that supports the methodology and is capable of being a complete source of information. The concept of CD is based on the division of information used in the systems of operational data processing (OLTP) and in the systems of receiving support. Such a subdivision allows optimizing both the structure of the operational storage data to perform the operations of introduction, modification, destruction, and retrieval, search, and so on.

There are two types of data sources:

1. Physical DS

The collected data are reduced to a single format, agreed and generalized. Analytical queries are sent to the data store.

Such a model leads to duplication of information in the DS. However, such a surplus does not exceed 1%. This is explained by the following reasons:

– when loading information in the DS, the data is filtered. many of them do not enter the DS, as they have no content from the point of view of use in the analysis procedures;

– generalized information is stored in the DS;

– during loading in the DS, the data is cleaned (unnecessary information is deleted) and reduced to a single format. After such processing, the data occupy a much smaller volume.

2. Virtual DS

The redundancy of information can be reduced to zero, using virtual DS [2; 3; 9]. The main advantages of virtual DS are:

– minimization of the amount of memory occupied by data on the information carrier;

– work with current, detailed data.

However, such an approach has its drawbacks. The processing time of queries to the virtual data warehouse significantly exceeds the corresponding indicators for the physical storage. In addition, the structures of the operational databases that are required for intensive data updates are highly normalized. However, in order to perform an analytical request, it is required to combine a large number of tables, which also leads to a decrease in speed. An integrated view of the virtual environment is possible only if the conditions of permanent accessibility of all operational data sources are met. Thus, the temporary inaccessibility of at least one of the sources can lead either to non-fulfillment of analytical queries or to inaccurate results. Execution of complex analytical queries takes up a large amount of computer resources on which they work. This leads to a decrease in the speed of OLTP-systems, which is unacceptable, since the time of performing operations in such systems is a very critical parameter.

The main disadvantage of the virtual data storage is considered to be the practical impossibility of obtaining data for a long period of time. In the absence of a physical storage, only the data available at the time of the request are placed in the operational data sources. As the data "ages", they are unloaded into applications and removed from the operational database [2; 11; 12].

Main characteristics of DS:

– Problem-oriented orientation. The data are grouped into categories and stored according to the areas they describe, not with the appendices they use. The information in the data warehouse is organized in accordance with the main aspects of the enterprise's activity (customers, sales, warehouse, etc.). This distinguishes the data warehouse from the operational database, where the data is organized according to the processes (extract of packages, unloading of goods, etc.). Substantial organization of data in the warehouse helps to significantly simplify the analysis, as well as increase the speed of execution of analytical queries. It is expressed, in particular, in the use of others, compared with the operating systems, the system of data organization [3; 7; 8; 10].

– Integration. The data are combined in such a way that they meet all the requirements of the enterprise as a whole, and not a single function. Before entering the data warehouse, the operational data is checked, cleaned and in a certain way aggregated. The output data is extracted from operative databases, transferred, cleaned, reduced to a single species, aggregated and collected in a similar way. It is much easier to analyze such integrated data.

– Incoherence. The data in the data warehouse are not created: that is, they come from external sources, are not corrected and are not deleted. Once trapped in a certain "historical hat", the data will never change. The stability of the data also facilitates their analysis.

– Dependence on time. The data in the storage are accurate and correct only in the case when they are tied to a certain interval or moment of time. The data in the storage are always directly related to a certain period of time. The data obtained from operational databases accumulate in the storage in the form of "historical layers", each of which consists of a specific period of time [6, 8, 11].

Historically, data warehouses are divided into the following three types (from the point of view of using models for data presentation):

– pelation;
– rich rich with the use of a universal rich model;
– hybrid.

Data storage facilities are better adapted to the storage and analytical processing of large volumes of data and, in the main, are the integration of pelation and data. They have the means to integrate data from different sources and allow you to work with both detailed and aggregated information. Therefore, as a rule, today the owners' own data are stored in pelationary databases [3; 11; 12].

The advantage of a pelation database is that it achieves a higher level of data abstraction, so data storage will be the most accurate.

Data flow diagram

A data flow diagram (DFD) is a model that represents a simulated system as a network of interconnected jobs. They can be used as additions to other models (such as IDEF) for visual display of current operations of document circulation in computing and computing (fig. 1).

Heine-Capcon notation is most often used in DFD to represent all elements. The system model in the context of DFD is presented in the form of some information model, the main components of which are the late data streams, which carry and transfer. Each of the subsystems performs certain transformations of the incoming data stream and transmits the results of the information processing in the form of data streams for other data. The main components of the data flow diagrams are as follows:

– external beings;
– processes;
– data flows;
– data accum. or storage facilities.

The data storage device can be physically realized by various means, but most often its realization in an electronic form on magnets is provided.

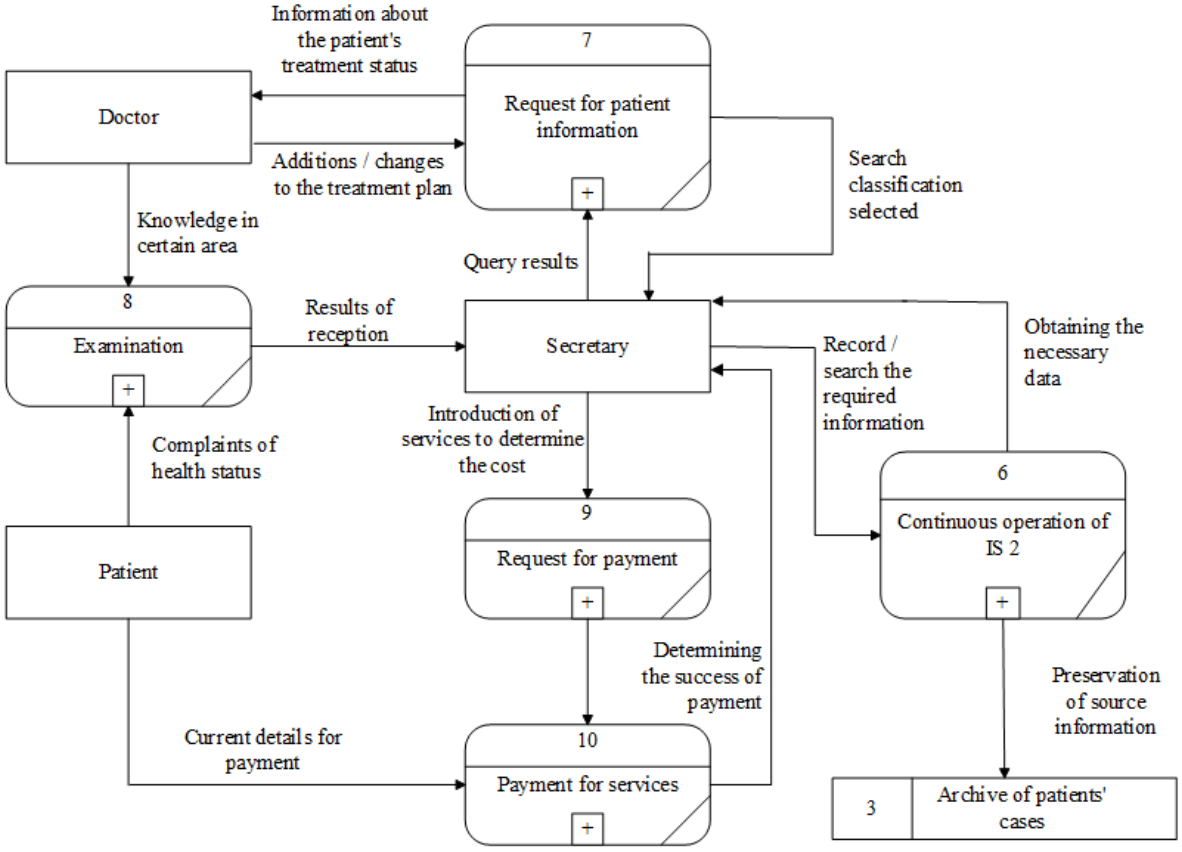

Figure 1 – Example of a DFD chart

On DFD diagrams, the data storage is represented by a rectangle with two fields; according to the rules, they are represented by a narrow rectangle open on the right side (Fig. 2).

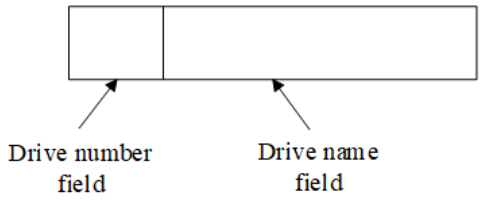

Figure 2 – Image of the drive structure on the data flow chart

The first field is used to indicate the name or identifier of the drive, and the second field is used to indicate the name. In this case, it is recommended to use a noun as the name of the drive, which characterizes the way of storing the relevant information [3; 12; 15].

Finally, the flow of data defines a qualitative feature of the information that is transmitted through some connection from the source to the receiver. The real data flow can be transmitted between two computers or any other method that allows data retrieval and recovery.

Internal representation of the DS

The data warehouse is an abstract device for storing information, which can be placed in the storage at any time, and after some time it can be taken away. This is an internal data warehouse for processes in the system. The data received before the treatment and the result after the treatment, as well as the intermediate values that must be preserved. This is the point: databases, tables or any other version of the organization and storage of data. Characteristics of data warehouses and external objects must be listed in the data dictionary, which is an internal database, which contains centralized information on all types of data, their names, structure, as well as information on their use [5; 8].

The advantage of the data dictionary is in the effective accumulation and management of information files of the subject area. Its application allows to reduce the superiority and contradiction of the data during their introduction, to carry out simple and effective management during their modification, Understand the database design procedure for data management centralization, establish links with other users. Thus, the data dictionary contains a generalized representation of all three levels: conceptual, logical, and physical, as well as information about the volume.

It should be noted that the data warehouse is used to execute unknown queries in advance. It can be physically implemented in the form of a drawer in the cabinet, a table in RAM, a file on a magnetic notion, etc. The name of the store is chosen from the understanding of the greatest information for the designer.

Features of construction and structure of DS

On the left in the rectangle there is a zone where the identifier of the data warehouse is located. However, there is a zone of the name, which briefly characterizes the set of elements of information support of the system, which are stored in the present day. The figure shows an example of the designation of data stores on DFD diagrams.

It is necessary to specify the identifiers:

– capital letters “D” or “C” (from the English words “digital” and “computer”) should be used to determine the computer form of data storage;

– storage facilities for the temporary storage of data that disappear after processing, should be marked with the letter "T" (from the word "transient");

– The places that determine the places of storage of the information presented in the traditional paper forms are identified by the letter “M” (from the manual).

So, the letter portion of the data warehouse identifiers determines their type. However, identifiers also have a numerical part. Let us consider the peculiarities of the formation of this part of the identifiers.

For DFD 1st level, this part should be in the form of a decimal digital suffix, which will define these nodes as the main ones for the whole data flow model.

The left side of this rectangle can have a double line to indicate the need for duplication of any joint on a specific DFD (Fig. 5).

Figure 5 – Duplicate data store

For diagrams of lower levels this is not enough, therefore identifiers of stocks of lower levels after the letter have a process number, an element of decomposition of which is present.

It should be noted that the DFD of the lower levels of hierarchy is the result of decomposition of the processes of the higher levels of of the hierarchy. Thus, if the process of the 1st level can be postponed, then it is the reason for the construction of the DFD of the 2nd level and so on. If there is no special need for this, there is no point in building a DFD below the 3rd level of the process.

For example, the decomposition of the process in DFD 1st level will mean, that the numerical part of the identifiers of all data stores located on its DFD of the 2nd level will begin with the number "6" followed by the symbol "/", and then the serial number of the compound; and so on for all levels of development of processes. This is illustrated below (Fig. 6):

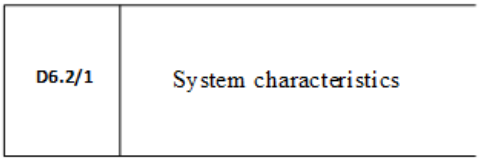

Figure 6 – Lower level

To display the results of decomposition of data warehouses from DFDs of higher levels on diagrams of lower levels, the identifiers of these storage units can be encoded with a letter. At the same time it is necessary to use small letters of the Latin alphabet which add to the basic identifier. For example, the decomposition elements of the D6 storage can be identified as D6a, D6b, D6c,…, and for D6 / 3 it will be D6 / 3a, D6 / 3c.

In material systems, data warehouses are displayed where objects are waiting to be processed, for example, in a queue. In data processing systems, data storage is a mechanism that allows data to be saved for future processes. In contrast to the arrows that describe objects in the hand, data warehouses serve to describe data that are temporarily not used, are found in mucus.

The data warehouse allows to define the data which will be stored in memory between processes on certain sites. In fact, the phenomenon represents the " slices" of data flows in time. The information it contains may be used at any time for its definition, and may be selected in any order. In general, the data warehouse is a preview of the future database, and the description of the data stored on it must be related to the information method.

Entity-relationship model or entity-relationship diagram (ER-model) is a data model that allows you to describe conceptual schemes with the help of node. ER-model is a meta-model of data, that is, a description of data models. ERD is designed to support data models and provides a standard way to define data and the relationship between them. In fact, with the help of ERD the detailing of data warehouses of the designed system is carried out [7 ... 12, 16].

In the event that the flow of data enters the warehouse or leaves it and its structure corresponds

According to the structure, it must have the same name and there is no need to reflect it on the diagrams.

Interrelation of SD with other components

As an example, let's look at a simplified model of the process of enrolling a patient in a hospital.

The external entities for this case are the patient and the secretary who registers the patient.

The data accumulator can be a database of documents and information created during the treatment of a patient who no longer needs treatment. That is, saving more irrelevant information in the archive.

Some data streams reflect the nature of the information needed to serve the hospital client. The corresponding model for this example can be observed in the form of diagrams of data streams (Fig. 8).

At the time, when data streams feed objects in the process of moving them, data warehouses model them in all other states. During the modeling of the production system, the data warehouses serve as temporary storage places, where the products are stored at the intermediate stages of processing.

In information systems, data storage mechanisms provide mechanisms that support data storage for their intermediate processing [4, 17].

For the correct representation of all elements by diagrams, there are general rules that are valid for DFD of all levels of packing. These rules are illustrated as a relationship matrix in [3].

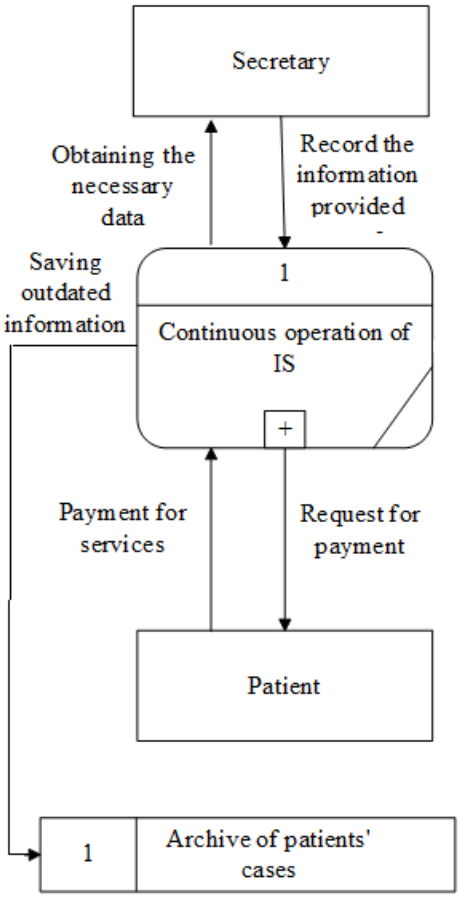

Figure 8 – Simplified diagram of data flows

It is known that the physical model reflects the parameters of the specific data points, where the data are stored in accordance with the logical model, as well as the function. The structure of this model may be different depending on the data notions and the processes of interaction between the notions in the storage.

The logical model reflects the structure of the data (data elements and the connections between them), which is available from the data carriers of the request to perform requests to them. To obtain data through this toolkit, a mapping must be constructed between conceptual and logical models.

When the procedure of transition from the physical level of reflection of properties and characteristics of the system to the logical representation is performed, the performance is performed. This is done by looking at the features of the logical model of the given system, which in this case is the guide and the means of checking the distribution of information. The following facts must be confirmed:

– each such site stores one or more information objects;

– there are no information objects that are stored in more than one data warehouse.

The terms of the urgent data are included in the corresponding data streams and thus are allocated from DFD. The correctness of connections of processes and data centers is also checked.

Functional and components of SD

Functional data warehouse architecture contains the following components:

– data center;

– client part of the system (warehouse designers, application development tools, administration tools, data analysis tools;

– loading the metadata dictionary from the XML file into the repository and exporting it from the repository to the XML file;

– Data Exchange Server – a set of data import / export programs from the repository and directories for organizing data exchange with external OLTP systems;

– Libraries of previous classes: ACL (Application Class Library), VCL (Visual Component Library), Win Lite [7; 14].

Filling of information stores takes place in several stages:

– extraction (extract) – import of data into the warehouse with information subdivisions, production departments and other sources;

– transformation – consolidation, aggregation of data, beating them into fractions, corrections and transformations into appropriate formats;

– download – to storage, synchronization with date or external events.

Maintenance of information repositories consists in: copying of databases, settings, typing, sending of fixed databases to the archive, management of users' rights, creation and editing of graphical diagrams of databases, etc.

The most important element of CD is the semantic shap – a mechanism that allows analytics to manipulate data with the help of business terms of the subject area. The semantic cap gives the user the opportunity to focus on the analysis and not to think about the mechanisms of data retrieval.

In addition to some data, the data warehouse contains other components that provide it with comprehensive work. Let's look at these components.

1. DS participants

All students are divided into three main groups:

a) end users;

b) development group;

c) support group.

As the technology of data storage is implemented through the gradual introduction of information use projects in certain problem areas, it is possible to learn about them.

End users (analysts; middle and top management). The data for all levels of aggregation-analytics are used most intensively. Their task includes a deep and thorough study of the data with the use of all available means of analysis. Using tools such as systems to support the decision-making of analysts in the process of research, they look for regularities between the data and present the results in the study.

The middle link uses DS data for information support of writing.

Top management uses highly aggregated data on key indicators that reflect the activities of the organization as a whole to make strategic decisions, using special applications in the form of interactive reports.

Development group (task setter; data designer; system support; setup guide). The director of tasks investigates information needs of the organization, 33 accessible sources of information, allocating problematic areas of activity of the organization and form. Close interaction with the end users and the CD administrator is a necessary requirement for the task setter. Only in this case, the end users will have access to the information they need with the help of the most effective tools.

The main task of the data designer is to create a logical structure of the DS, which provides effective access to all necessary data.

Systematic support of DS provides definition and writing of technical problems of creation and development of DS. This includes the choice of platform, OS, the necessary requirements for memory and disk space, data security, etc.

When researching the subject area, it may appear that a set of standard measures is not enough to use the information stored. In this case, the user guide deals with the implementation of specialized analysis tools.

Support group (DS administrator; user support). The functions of the DS administrator differ significantly from the functions of the database administrator. The functions of the DS administrator are focused on maintaining the quality of the data of the repository, while the functions of the database administrator are focused on the technical support of the database. The main instruments of the DS administrator are the loading monitor and the data usage monitor.

By controlling the data loading information in the DS, the DS administrator monitors the execution of the data loading procedure, as well as the information of the data loading. The main instruments of the DS administrator are the loading monitor and the data usage monitor.

By controlling the data loading information in the DS, the DS administrator monitors the execution of the data loading procedure, as well as the information of the data loading violations, communicates with the persons responsible for the source. The heads of the subdivisions that collect the primary data are usually responsible for the source of data. The information about the use of data in the DS is used by the administrator when interacting with the development group to increase the efficiency and reliability of work. The support of end users provides for the development of methods and recommendations for the use of analysis tools, training of specialists, and training.

Software (loading tools; monitoring tools; creation and development tools). To perform data loading in the DS, the following software is used: process manager; data loader; data analyzer.

1. Process dispatcher. Regular loading of new data is necessary for normal functioning of the DS. Observation of the procedure for loading data from the source is a necessary part of building a logical structure of DS. The dispatcher carries out loading procedures in accordance with the regulations.

2. Data loader. Data sources for DS can be the most well-known databases, as well as external data in other formats. It is important that the data loader has access to the maximum number of DBMSs and other data formats. The function of the loader also includes the transformation of data into the specified format.

3. Data analyzer. The quality of the data, taken from different sources, often leaves much to be desired. Automatic analysis of data for accuracy and inconsistency is an important part of DS technology.

Monitoring tools (load monitor, data usage monitor).

Download monitor. Purpose of use DS technology – providing reliable data for analysis tools, so the necessary tool is a data download monitor that collects information about the execution of data download processes and informs the administrator about the progress of these processes.

Data usage monitor. The data usage monitor is a very useful component of the software to increase the efficiency of data access, and can also provide information about the possibility of translating the lost detailed data into the status of obsolete detailed data [9; 11; 13].

Means of creation and development (DBMS of data warehousing; Means of data structure management; Means of data source management; Means of construction of specialized databases).

DBMS data storage should be focused on the features of DS technology – to work with large volumes of data, to ensure the necessary security to allow to create very difficult data structures (such as rich databases), to carry out fast, packed on many users, to do.

Means of managing the structure of data DS. For fast realization of logical structure of the data it is necessary to have a convenient interactive means of management of structure of DS. The quality of this tool determines the speed of development and development of DS, so it is a very important factor.

Means of task of data sources.

They are used to create data sources that are loaded in the warehouse: to determine the connection between DS structures and sources, to create a process of transformation, purification.

For the construction of specialized data centers. Specialized data warehouses perform the functions of certain divisions and are an important part of DS technology, and with the development of DS It is often necessary to transfer specialized DSs to other technical equipment, so the tool must have a flexible interface to work with specialized DS, [10…16; 18].

Requests to DS

DS is intended for certain types of queries. Since one of the main functions of the data warehouse is data analysis, and not their operational operation, the queries to the DS have a different character than the queries to the database [11; 12].

Slice-and-dice requests. The concept of the data is based on two main ideas: making a choice, reducing a cube. Fixation of the value of the measurements reduces the size of the cube, but at the same time more general operations are possible.

Drill-down and roll-up queries are reciprocal operations that use a set of measurements and parameters for aggregation. Generalization to higher values corresponds to the exclusion of pomp. For example, the summary from the "Misto" to the "Kraina" rows aggregates the meaning for Lviv and Kyiv in one meaning – Ukraine.

Drill-across queries combine cubes that have one or more common dimensions. From the point of view of pelationary algebra, such an operation performs a join. Queries of the form ranking returns only those results that appear in the upper or lower part of the list ordered in a certain way.

Remarks. It is not necessary to try to convey a geometric interpretation of the concept of "rich cube", as it is simply a service term, which describes the method of presentation.

Subspecies of DS

One of the subspecies of DS is the data source (DS) – from the data warehouse, the array of thematic, narrowly focused information, which is opioid, written. According to the definition of authors, the data sources are a set of thematic databases that contain information related to certain aspects of the organization's activities.

The next subspecies – operational data storage (ODS) – is a subject-oriented, integrated, variable set of consolidated data, which contains the current flow.

Like ODS, the data warehouse is a subject-oriented integrated set of consolidated data. From this point of view, they exist, as in both cases, and in another case, the data must be loaded from the transactional system. But this is where their life ends. ODS contains data that change, while in the data storage due to loading they do not change [15… 18].

Another difference is that the operating room contains only data that is current at a certain point in time, while in the store it is stored as current and current. At the same time, the actuality of data in the field is much lower than in the operational field. As a rule, data stored for the last 24 hours are stored in the storage room, while the actuality of data in the ODS can be measured in seconds.

Another difference between the ODS and the storage is that it contains only detailed data, while the storage contains both detailed and aggregated data.

Conclusion

In the last decade, DS technology has been rapidly evolving. Many companies offer DS writing in the marketplace, and thousands of organizations are already using this powerful tool to support analytical projects. If the DFD diagrams do not show the characteristics of the time of execution of certain processes and the transfer of data between the processes, then the system models, which are the same. Unfortunately, all these features of the methodology of structural system analysis limit the possibilities of its wide application. The main problem with the use of data warehouses in Ukrainian enterprises is that the ready-made business model supported by Western software products may not be suitable for Ukrainian systems. Domestic institutions do not have such essences that are in demand in foreign systems. The second, no less important difficulty, from the point of view of use of products of such class, is their high price. The data warehouse ensures the achievement of the following goals: ease of access, consistency of information, resistance to change, protection, support.
The spectrum of data storage technology is quite wide. The tasks written with the help of data storage are, as a rule, related to the tasks of management analysis and strategic planning.

References

1. Data warehouse. – https://uk-wiki.ru/wiki/%D0%A1%D1%85%D0%BE%D0.
2. Wikipedia. data warehouses. – https://uk.wikipedia.org/wiki/%D0%A1%D1%85% D0% BE% D0% B2% D0% B8% D1% 89% D0% B5_% D0% B4% D0% B0% D0% BD% D0% B8% D1% 85.
3.Tsiutsiura, S. V., Wart, E. V. (2015). Methods of projection of object models on data structures. *Management of complex systems development*, 21, 91–99.
4. Chernyshev, D. O., Kyiv, K. I., Tsiutsiura, S. V., Tsiutsiura, M. I., Gots, V. V. (2019). Introduction of technology of modeling of information objects at stages of a life cycle. *Management of complex systems development*, 40, 140–146.
5. Fedusenko, O. V., Tsiutsiura, M. I., Fedusenko, A. O., Tsiutsiura, S. V. (2016). Development of an adaptive knowledge control system with open questions. *Management of complex systems development*, 28, 129–135.
6. Tsiutsiura, S. V., Krivoruchko, O. B., Tsiutsiura, M. I. (2012). Key performance indicators. Principles of development of key efficiency indicators for the budget sphere. *Management of complex systems development*, 10, 87–92.
7. Tsiutsiura, S. V., Krivoruchko, O. V., Tsiutsiura, M. I. (2012). Theoretical foundations and essence of management decisions. Models of management decision making. *Management of complex systems development*, 9, 30–38.
8. Kozlovskaya, V. P. (2016). *Methods and means of data integration*, 31–35, 56–58.
9. Method of modeling data flows. https: //studfile.net/preview/3030895/page: 3 /
10. Methodology of using databases. https://stud.com.ua/35675/informatika/metodologiya_vikoristannya_danih.
11. Erukaev, A. V., Fedusenko, O. V. (2016). Application of production rules for the implementation of genetic algorithm. *Science and Education a New Dimension. Natural and Technical Sciences,* IV (12), 110, 31–33.
12. Yatsyshyn, A. Yu. (2006). Information technology of construction of distributed storage of hybrid data, 90.
13. Grigoriev, Yu. A., Revunkov, G. I. (2002). Data banks. Moscow : Izd-vo MSTU im. NE Bauman, 320.
14. Markov, A. (1997). The concept of building an electronic archive. *Open Systems*, 1, 54–58.
15. Krenke, D. (2003). Theory and practice of building databases. St. Petersburg : Peter, 800.
16. Zavadsky, I.O. (2011). Basics of data: Kyiv: Publisher I. O. Zavadskii, 192.
17. Pushnikov, A. Yu. (1999). Introduction to database management systems. Part 1. Pelation model of data: training manual. Bashkircko University Press. Ufa, 108.
18. Pasichnyk, V. V., (2006). Shakhovska NB Data Warehouses: Tutorial. Lviv: Magnolia, 496.

Received 14.10.2020

Чернишев Денис Олегович
Доктор технічних наук, професор, перший проректор, *orcid.org/0000-0002-1946-9242*
Київський національний університет будівництва і архітектури, Київ
Цюцюра Світлана Володимирівна
Доктор технічних наук, професор, заведувач кафедри інформаційних технологій, *orcid.org/0000-0002-4270-7405*
Київський національний університет будівництва і архітектури, Київ
Лященко Тамара Олексіївна
Старший викладач кафедри інформаційних технологій, *orcid.org/0000-0001-9092-0297*
Київський національний університет будівництва і архітектури, Київ
Лузіна Юлія Володимирівна
Студентка IV курсу кафедри інформаційних технологій, *orcid.org/0000-0002-8426-9431*
Київський національний університет будівництва і архітектури, Київ
Бородиня Віталій Віталійович
Аспірант кафедри інформаційних технологій проектування та прикладної математики, *orcid.org/0000-0001-9513-92645*
Київський національний університет будівництва і архітектури, Київ

ЗАСТОСУВАННЯ СХОВИЩ ДАНИХ НА ПРИКЛАДІ СИСТЕМИ ОБРОБКИ ЗАЯВ У ДЕРЖАВНИХ УСТАНОВАХ

Анотація. *У статті висвітлено питання про важливість інформації, необхідність забезпечення її відповідного зберігання та використання. Нині постає питання важливості вивчення функціонування засобів збереження даних у одній з найпоширеніших моделей проєктування – Data Flow Diagram, та сховищ даних (СД) в цілому. Проведено дослідження абстрактних пристроїв цієї моделі для збереження інформації: різновиди, особливості та види сховищ даних, які використовуються в інформаційних системах (реляційні, багатовимірні та гібридні сховища), з точки зору використання моделей для представлення даних; зовнішній вигляд сховищ та правила побудови; літерні ідентифікатори сховищ: "D", "C", "M" та "T", за допомогою яких визначається тип сховищ; особливості числової частини ідентифікаторів для декомпозицій процесів першого та другого рівнів; механізми, що підтримують зберігання даних для їхньої проміжної обробки в інформаційних системах. Перехід відображення властивостей та характеристик від фізичного до логічного представлення, раціоналізація сховищ даних шляхом розгляду особливостей логічної моделі. У процесі роботи розглянуто особливості побудови DFD і відображення взаємозв'язків усіх комплектуючих діаграм, визначені у загальних правилах, що є дійсними для DFD усіх рівнів ієрархії. Висвітлені питання стосуються елементів діаграм, які можуть бути вільно експлуатовані в межах України, адже елементне наповнення різних типів діаграм істотно відрізняється залежно від країни-розробника. Своєрідний функціонал вітчизняних DFD-діаграм дає змогу використовувати для побудови достатній, хоч і обмежений елементний склад. Залежно від фактичних можливостей, діаграми Українського виробника дещо спрощені та неудосконалені відповідно до сучасних технологій, але ці фактори не зменшують важливість і необхідність відповідного захисту даних на найвищому рівні. Тому в цій роботі висвітлено основні найбільш інформативні моменти стосовно використання того чи іншого виду сховищ, затребуваність СД у повсякденній роботі, можливі переваги та недоліки використання фізичних сховищ даних та особливості експлуатації віртуальних.*

Ключові слова: діаграма потоків даних; сховище даних; база даних; збереження інформації

Link to the article

APA Chernyshev, Denys, Tsiutsiura, Svitlana, Lyashchenko, Tamara, Luzina, Yuliia, & Borodynia, Vitalii, (2020). The use of data warehouses on the example of the application processing system in government agencies. *Management of Development of Complex Systems*, 44, 166 – 174, dx.doi.org\10.32347/2412-9933.2020.44.166-174.

ДСТУ Чернишев Д. О., Цюцюра С. В., Лященко Т. О., Лузіна Ю. В., Бородиня В. В. Застосування сховищ даних на прикладі системи обробки заяв у державних установах. *Управління розвитком складних систем*. Київ, 2020. № 44. С. 166 – 174, dx.doi.org\10.32347/2412-9933.2020.44.166-174.