Virtual geoarchaeological repository of Baikal Siberia

Fereferov E.S.[1[0000-0002-7316-444X]], Vetrov A.A.[2[1111-2222-3333-4444]] and Berdnikova N.E.[2[0000-0002-2470-6230]]

[1] Matrosov Institute for System Dynamics and Control Theory of SB RAS, 664033, 134 Lermontov St., Irkutsk, Russia
[2] Scientific Research Center «Baikal Region», Irkutsk State University, 664003, 1 Karl Marx St., Irkutsk, Russia
fereferov@icc.ru

Abstract. The article discusses the problem of creating an interdisciplinary database to support research on the project "Baikal Siberia in the Stone Age: at the crossroads of worlds." The architecture of an integrated information system for supporting archaeological research for the formation of a database has been developed. The information system is deployed on two servers that solve separate tasks and is supported by a network storage system. The subsystems are developed using the technology of creating information systems based on the specifications of database applications. Specifications contain, in pure form, the minimum necessary information about database tables, their fields, relationships between them, and their use in the database application. This approach allows you to easily and quickly modernize the information system throughout the entire life cycle. The creation of databases and application specifications is based on the developed standards for geoarchaeological research. In addition, tools were developed to solve specific problems of archeology. For example, tools for creating separate portable subsystems for working in the excavation and publishing 3D models.

Keywords: Information System, Declarative Specification, GIS, Archaeology, 3D-modeling.

1 Introduction

At present, archaeological information on the Stone Age of Baikal Siberia is presented in the form of diverse, poorly organized and poorly structured data. This is subject material, text and graphic documentation, photo and video recordings, geospatial, analytical data. The information is kept in various regional, state institutions, museums, educational institutions, archives, state bodies for the protection of cultural heritage, private companies, including individuals.
This situation restricts access to the entire volume of data, does not allow establishing relationships between them and drawing meaningful conclusions based on them. The development and creation of a system for processing large amounts of data for the

Stone Age of Baikal Siberia is the first experience of a synthetic database in Russian archeology and in the future can become the core for the formation of Big Data on the evolution of human and societies at the regional and global levels.

2 Related work

The task of structuring archaeological data in order to search and analyze information has existed since the emergence of archeology as a science. Large volumes of archaeological research results have been accumulated in the form of printed documents. The digitization of old archaeological data, their integration and the creation of associated databases are important tasks. At present, a great deal of experience has been gained in the formation of archaeological databases. A variety of digital data formats are used to store the characteristics of archaeological sites: spreadsheets, documents of various word processors, 2D and 3D images, CAD files and digital maps. Database management systems (DBMS) are one of the base tools for the formation and maintenance of archaeological databases.

The development of information systems (IS) for solving archeological problems has long been carried out in Russia and in other countries. Due to its specificity, archaeological research is closely related to the use of spatial data. Therefore, support systems for archaeological research are often developed on the basis of geographic information systems (GIS).

Archaeological Information System (AIS) "Archeograph" [1] was developed in 2008 at the Institute of the History of Material Culture of the Russian Academy of Sciences. It is one of the first Russian systems for describing archaeological sites. AIS "Archeograph" is implemented as a desktop system and provides accounting and storage of basic information about archaeological sites and artifacts. The system also supports interaction with GIS MapInfo through DDE technology (Dynamic Data Exchange). Today the project is not developing, and the software is morally outdated.

At the Institute of Archeology (IA) RAS developed GIS "Archaeological monuments of Russia" [2]. This GIS is positioned as a system of accounting for objects of archaeological heritage on a national scale. The system is a desktop software and provides the collection and integration of information on archaeological sites based on the processing of reports on archaeological excavations received in the scientific archive of the IA RAS. The system includes a module that allows you to save the results of user queries in the KML/KMZ file formats and then visualize it with well-known geo-services (for example, Google Maps, Yandex Maps).

The transition from traditional databases to the formation of Big Date systems in the practice of world archeology is gradually being implemented. For example, the project on the virtual repository of zooarcheological data of the Arctic [3] , the international repository of digital records of archaeological research[4], the national geographic information system "Archaeological Monuments of Russia" [5], the project “The Endangered Archeology in the Middle East and North Africa (EAMENA)” [6]. Each archaeological database is created for a specific purpose.

3 Georchaeological standards

We have divided the standards for geo-archaeological sites according to research fields. When forming the standards, we took into account the fact that a standard can have a long and short record. A long record is focused on recording the maximum amount of information on certain phenomena, essences of a geoarchaeological object. A short entry is a shortened version of a long entry. It can be used to make information publicly available or used for other purposes. Currently, we have developed 28 standards for recording information for various research fields.

The basic standard is a standard that reflects the general characteristics (external and generalized essence) of a geoarchaeological object. This standard consists of 21 items, which take into account the identified name of the object, its geographical characteristics and typological essence, the current state, generalized history of research, the nature of the culture-hosting deposits, the taphonomy of archaeological material, the characteristics of cultural complexes, a list of documentation and publications on the object.

This standard links all other standards from different research fields into a single structure. Most of the developed standards allow using them both in field research and digitalization of archival materials and data. We have developed standards for the spatial characteristics of the object; on the description of semi-structured and unstructured information; on the description and analysis of deposits, radiocarbon dating; on ground photographic recording; on the field fixation of archaeological material.

The block of standards for describing natural science data consists of standards for describing sediments, various samples, chemical, granulometric, paleomagnetic and other multi-proxy records, a standard for morphogenetic analysis of sediment structure. These standards establish unified algorithms for describing the structure of deposits of geoarchaeological and geological objects by using pedolithological methods (coupled methods of soil science and geology) [7], which allow describing deposits with greater detail and accuracy, as well as interpreting them in a single system. The standard for describing faunistic collections has been developed on the basis of standardized methods for describing and measuring faunistic remains, analysis of their taphonomic characteristics [8–11]. A standard for radiocarbon dating samples has also been developed. which reflects the characteristics of the sample, the results of dating and data on isotopes, collagen content, necessary for further analytical and interpretation operations.

Separately, we have developed standards for describing object material: stone, bone, and ceramics. When creating standards for artifacts, we were guided by the approaches and methods of descriptive analysis developed by J.C. Garden [12-13]. We also used the schemes, methods of formalized processing, description, classification, typology, proposed in Russian archeology [14]. When forming the standards, we relied on both the existing systems of description, classification, typology of stone artifacts [15–18]; bone tools [19, 20]; ceramic vessels [21]; and for the developments carried out at the Irkutsk University [22, 23].

We used the maximum number of qualitative and quantitative features to form long records of standards for artifacts. Standards for chips, cores, scrapers, ceramics, harpoons, jade products have been developed. Standards for other groups of artifacts are under development. All standards are accompanied by catalog-thesauri, which explain the concepts and present the terminology used.

4 Georchaeological repository architecture.

Geoarchaeological research is a complex process and requires the use of modern data processing methods, information technology (IT) and tools. Such researches are multi-stage and multi-layered with a high level of detail. The volume of information received often approach Big Data, which imposes increased requirements on the database management systems (DBMS) used and technical storage facilities. When choosing a DBMS, it is also necessary to take into account the diversity and characteristics of the research fields of geoarchaeological research. For example, there is a need to work with data in the field on mobile devices in the absence of access to servers. The formation of new standards for describing research and their development requires the use of flexible tools for creating and maintaining information systems throughout the entire life cycle.

Based on the above problems, we have proposed the following architecture for the functioning of the "Virtual geoarchaeological repository of Baikal Siberia" (see Fig. 1). There are two modes to interact with the repository: full-featured (closed) only for employees of the Research Center “Baikal Region” and common (presentation) for a wide range of researchers.

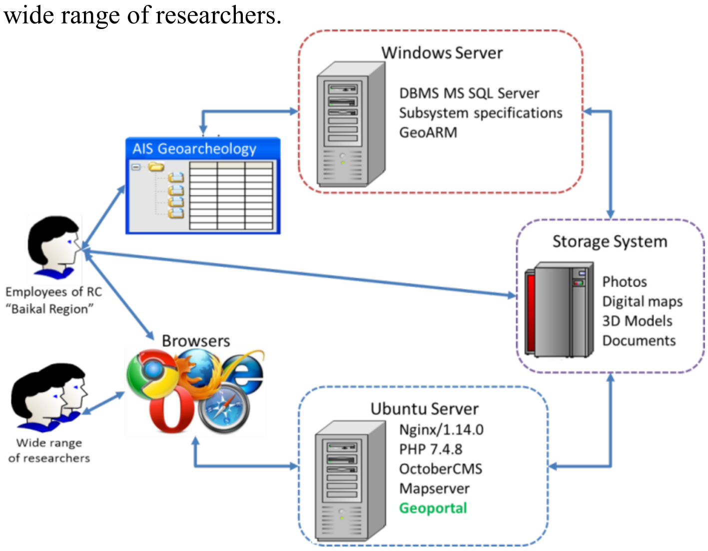

Fig. 1. The architecture of the repository functioning

The main full-featured part of the repository was installed on a server running Windows Server. We installed the MS SQL Server database management system on the server to provide interaction with the AIS Geoarcheology database. Also we have installed the specifications of the AIS "Geoarcheology" subsystems on this server.

We use the "GeoARM" tool system for creating and maintenance of subsystems specification. MS SQLServer provides storage and processing of big volumes of subject data and service information. Data storage reliability is achieved by mirroring databases and repository files to a network storage system (NAS). We have developed a portal to provide a wide range of users with access to research results. This portal is function on a server run by Ubuntu. We have implemented a portal based on October CMS. The NGINX web server and Mapserver were installed and configured on the server for the portal to function. The portal provides limited access to the AIS "Geoarcheology" database , allows you to filter data and display it on a digital map. The portal also provides access to photos and documents stored in the network storage.

We used our model-driven approach to increase the speed of creation and the flexibility of maintenance of the AIS Geoarcheology subsystems. In our approach [24,25], the application model is represented in the form of a declarative specification containing sufficient information to automatically create the user interface of the database application, ensure the execution of CRUD functions, build queries and interact with spatial data (PD). Representation of application models in the form of specifications allows us to support modular development of information systems: to integrate and combine ready-made application specifications when developing new ones. Ready specifications can be used to automatically create both desktop (Desktop) and web systems.

Automation of the development of subsystems is achieved through the use of the GeoARM tool system. This tool system provides the interactive formation of all elements of the database application models, as well as the dynamic creation of subsystems as a result of the interpretation of the created specifications. The developed standards for describing geoarchaeological studies play an important role in the development of database structures and the creation of subsystem specifications. Each standard contains a ready-made set of attributes, links between them and is the basis for the development of structures of database tables and models of specific subsystems of the AIS "Geoarcheology".

The AIS “Geoarcheology” created by us is a complex of interconnected flexible subsystems and thematic databases, provided with a reliable backup storage. The architecture of the AIS is shown in Figure 2. Based on the developed standards with the help of GeoARM, we have created the specifications of the subsystems responsible for certain aspects of the description of archaeological research. It is enough to correct only the specification when modernization a standard need (adding new attributes for example). The corresponding subsystem is generate automatically based on the interpretation of the specification. Our technology does not require making changes to the program code and recompiling the subsystem when changing the database structure.

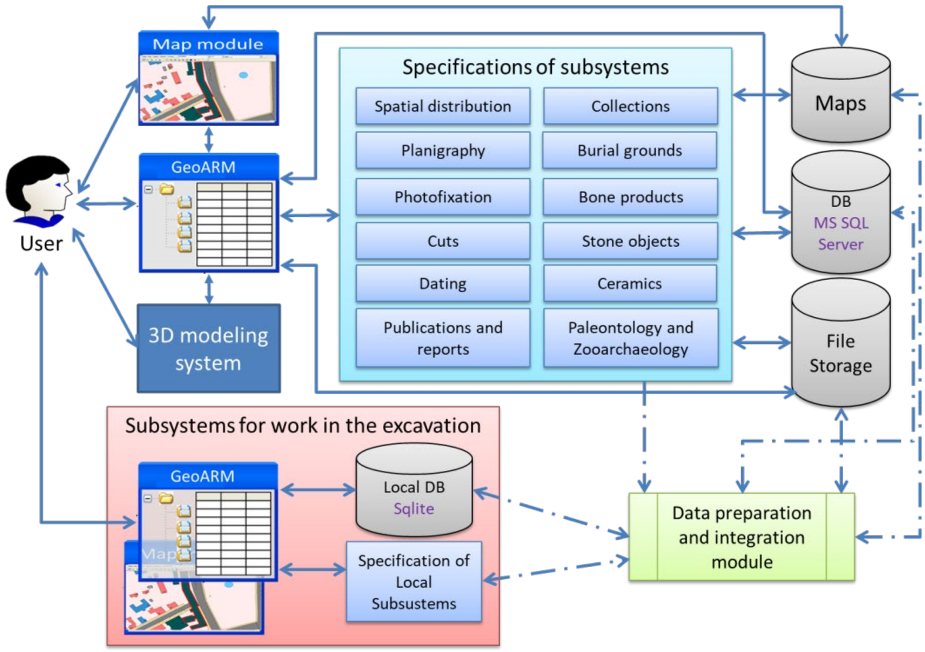

Fig. 2. Architecture of the AIS "Geoarcheology"

We have developed database structures and created specifications for the following subsystems: Geoarchaeological objects, Sections, Layers, Field fixation, Photo fixation diary, Archaeologists, Publications, Chips. The specifications of the individual subsystems are integrated into one common specification, allowing users to interact with both individual information and data across the entire archaeological site. The AIS "Geoarcheology" has a typical interface (see Fig. 3). Entity tree (subsystems, dictionaries) is located on the left of the program window. The content of the corresponding entities is displayed in the center to the right of the entity tree. Working with data is provided in two modes: in tabular mode and form mode with record fields. In form mode, the user can select values from dictionaries and interact with the records of the detail tables. Subsystem "Map" can be used to work with spatial data.

The AIS "Geoarcheology" provides users with access to thematic, cartographic, photographic databases, as well as tools for visualization and data analysis. The GeoARM instrumental system supports interaction with external subsystems, which allows you to expand the functionality of applications without modifying the code of the system itself. We have implemented support for storing and presenting 3D models of objects in the AIS Geoarcheology. Researchers obtain these models from 3D scanners or create based on photogrammetry methods in well-known modeling systems (for example, Agisoft, Meshmixer). Today, the system supports work with an open geometry description format - obj. We have implemented the visualization of 3D models of landscapes and archaeological finds by calling external systems. The

Tree.js library was used to visualize 3D models of objects. A 3D model of the Shishkinskaya Pisanitsa archaeological site has been created.

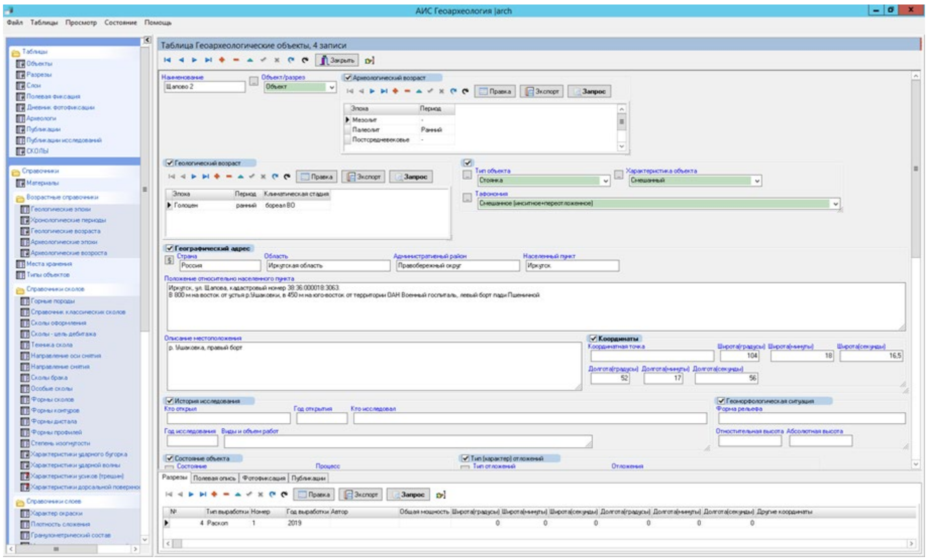

Fig. 3. Interface of AIS "Georcheology"

Often, archaeologists conduct research in places where there is no access to the Internet and, accordingly, to a common repository. We have developed a module for preparing and integrating data (MPID) to form subsystems that ensure work in such conditions. This tool provides generation of local off-line subsystems. With the help of MPID, the user can customize the configuration of the local subsystem by choosing the specifications of the required subsystems. Then the MPID will generate the necessary database tables for the SQLite DBMS and the subsystem specification for interacting with this DB. To work with subsystems in off-line mode, the researcher only needs to install a database file in SQLite format, the necessary specification and the GeoARM framework on his laptop. MPID allows loading the data collected using the off-line subsystem into the main database of the repository.

October CMS, which we used to create the portal, implements the MVC pattern. For automation, we have developed a translator of subsystem specifications into a set of files implementing the model in October CMS. The created portal (see Fig. 4) provides access to thematic and cartographic databases via a web interface. Authorized users have access to all information in the repository. Other users can only view the "short entry" about the archaeological sites. Cartographic layers can be created in the portal to display data on a map if archaeological objects have spatial attributes. The Mapserver we use generates images of map layers according to the WMS standard. We used the open source Leaflet library to implement the map display module (see Fig. 5). As a background, the portal uses well-known GIS maps, such as Google, YandexMaps, OpenStreetMap, 2GIS, Bing.

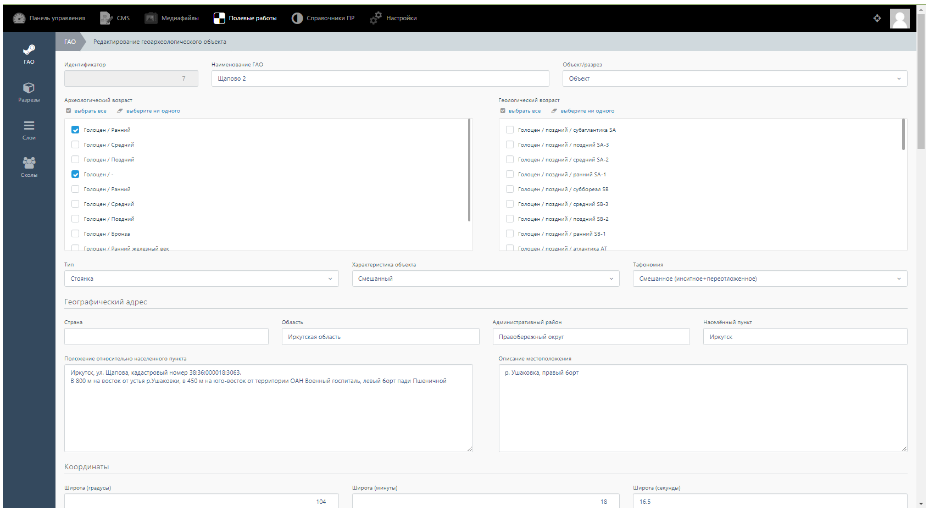

Fig. 4. The portal of geoarchaeological repository of Baikal Siberia

Fig. 5. The map display module

5 Conclusions

Archaeological research is almost always accompanied by a large amount of heterogeneous data: documents, photographs, maps, 3D models. The article deals with the urgent problem of creating an interdisciplinary database of the project "Baikal Siberia in the Stone Age: at the crossroads of worlds." The architecture of an integrated information system for supporting archaeological research has been developed, which ensures the formation of a database. The information system is deployed on two servers that solve separate tasks and is supported by a network storage system.

The subsystems are developed using the technology of creating information systems based on the specifications of database applications. Specifications contain, in pure form, the minimum necessary information about database tables, their fields, relationships between them, and their use in the database application. This approach allows you to easily and quickly modernize the information system throughout the entire life cycle. The creation of databases and application specifications is based on the developed standards for geoarchaeological research.

We also developed tools for solving specific problems of archeology. For example, a tool for creating local portable subsystems for working in excavations and publishing 3D models.

6 Acknowledgment

The work was carried out with financial support by the Government of the Russian Federation, project No. 075-15-2019-866 "Baikal Siberia in the Stone Age: at the crossroads of worlds". Results are achieved using the Centre of collective usage «Integrated information network of Irkutsk scientific educational complex».

References

1. Vasiliev, St.A.: AIS "Archeographer": a system for describing archaeological sites and data output in GIS. In: Archeology and computer technology: presentation and analysis of archaeological materials, pp. 13–21. Izhevsk: UdmIIAL Ural Branch of RAS (2005).
2. Makarov, N. A., Zelentsova, O. V., Korobov, D. S., Voroshilov, A. N., Chernikov, A. P. Geoinformation system “Archaeological Monuments of Russia”: methodological approaches to development and first results of filling. Brief Communications of the Institute Archeology, No. 237, pp. 7–20 (2015).
3. Virtual Zooarcheology of the Arctic Project (VZAP), http://vzap.iri.isu.edu/, last accessed 2020/08/11.
4. The Digital Archaeological Record (tDAR), https://www.tdar.org/, last accessed 2020/07/17.
5. Makarov, N.A., Zelentsova, O.V., Korobov, D.S., Chernikov, A.P., Voroshilov, A.N.: The first steps to create a national geographic information system "Archaeological sites of Russia". Archeology, ethnography and anthropology of Eurasia, Vol. 43, No. 4, 85–93 (2015).
6. Zerbini, A.: Developing a Heritage Database for the Middle East and North Africa. Journal of Field Archaeology, Vol. 43, Sup. 1, pp. 9–18 (2018).
7. Vorobyova, G.A.: Soil as a chronicle of natural events in the Baikal region (problems of evolution and classification of soils). Irkutsk State University, Irkutsk (2010).
8. Antipina, E.E.: Modern archaeozoology: tasks and research methods. Interdisciplinary integration in archeology (based on lectures for graduate students and young employees), pp. 96-117 (2016).
9. Baryshnikov, G.F.: Fauna of Russia and neighboring countries. Mammals. Bear family (Carnivora, Ursidae) Vol. 1, Issue 5. pp 542. Nauka, Saint Petersburg (2007).

10. Erokhin, N.G., Bachura, O.P.: A new approach to computerized formalization of fragmentation of mammalian bone remains in archaeozoological research. Methods of interdisciplinary archaeological research, pp 62–69. Nauka, Omsk (2011).
11. Van der Made J.: The rhinos from the Middle Pleistocene of Neumark-Nord (Saxony-Anhalt). Veröffentlichungen des Landesamtes für Denkmalpflege und Archäologie Vol. 62, pp. 433-527 (2010).
12. Gardin, J.-C. Four codes for the description of artifacts, an essay in archaeological technique and theory. American Anthropologist Vol. 60, No. 2, pp. 335–357 (1958).
13. Gardin, J.-C.: Methods for the descriptive analysis of archaeological materials. American Antiquity Vol. 32, Is. 1, pp. 13—30 (1967).
14. Kamenetskiy, I.S., Marshak, B.I., Sher, Ya.A.: Analysis of archaeological sources. Possibilities of a formalized approach. IA RAN, Moscow (2013).
15. Kolobova, K.A.: Techniques for the design of stone tools in the Paleolithic industries of Gorny Altai. Publishing house IAET, Novosibirsk (2006).
16. Nekhoroshev, P.E.: Technological method for studying the primary cleavage of the Middle Paleolithic stone. European House, Sant Petersburg (1999).
17. Pavlenok, G.D., Pavlenok, K.K.: The spinning technique in the Stone Age: a review of English and Russian-language literature. Bulletin of Novosibirsk State University, Series: history, philology, Vol. 13, Issue. 5: Archeology and Ethnography, pp. 26–37 (2014).
18. Shea, J.J.: Stone tools in the Paleolithic and Neolithic near East: a guide. Cambridge University Press, Cambridge (2013).
19. Clodoré-Tissot, T., Le Gonidec, M.-B., Ramseyer, D., Anderes, C. Instruments sonores du Néolithique à l'aube de l'Antiquité. Fiches de la Commission de nomenclature sur l'industrie de l'os préhistorique. Cahier XII, 89 (2009).
20. Mons, L., Péan, S., Pigeaud, R. Matières d'art: Représentations préhistoriques et supports osseux, relations et contraintes. Fiches de la Commission de nomenclature sur l'industrie de l'os préhistorique. Cahier XIII, p. 286 (2014).
21. Nikolaenko, S.N.: Description, comparison and differentiation of ceramic vessels in the Western Baikal region. News of the Laboratory of Ancient Technologies, Issue. 4, pp. 34–64 (2006).
22. Berdnikova, N.E.: Stone splitting and chip morphology. News of the Ancient technology laboratories, Issue. 1, pp. 58-66 (2003).
23. Lipnina, E.A., Lokhov, D.N., Medvedev, G.I.: About stone axes "with ears" - trunnion axes of Northern Asia. Bulletin of Irkutsk State University. Series "Geoarcheology. Ethnology. Anthropology", No. 1 (2), pp. 71-101 (2013).
24. Bychkov, I.V., Hmelnov, A.E., Fereferov, E.S., Rugnikov, G.M. and Gachenko, A.S.: Methods and Tools for Automation of Development of Information Systems Using Specifications of Database Applications. In: 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC), pp. 1-6. Vladivostok (2018)
25. Hmelnov A., Fereferov E. Development of cross-platform problem-oriented systems using specifications of database applications // CEUR Workshop Proceedings: Proc. of 2nd Scientific-Practical Workshop Information Technologies: Algorithms, Models, Systems (ITAMS'2019). 2019. Vol. 2463. pp. 59-69.