

DEVELOPMENT OF A SERVICE FOR CONDUCTING RADIOBIOLOGICAL STUDIES ON THE HYBRILIT PLATFORM

Yu.A. Butenko¹, D.M. Marov^{1,2}, A.V. Nechaevskiy¹, D.V. Podgainy¹

¹ *Laboratory of Information Technologies, Joint Institute for Nuclear Research, Dubna, Russia*

² *Federal State-Funded Educational Institution of Higher Education of Moscow Region “Dubna University”, Dubna, Russia*

E-mail: ybutenko@jinr.ru

The article is devoted to the development of an information system for analyzing experimental data on behavioral and pathomorphological changes in the central nervous system in the study of the effects of ionizing radiation and other factors on laboratory animals. Based on the analysis of similar existing systems, their major capacities and disadvantages are defined, and a technology stack for developing a prototype is determined. For data acquisition and storage, a client-server architecture of the system is elaborated, and a database is designed. The prototype of the user web interface, the main components and tools for analyzing experimental data are presented. The system is built on the basis of the HybriLIT heterogeneous platform.

Keywords: web service, web interface, client-server application

Yuriy Butenko, Dmitriy Marov, Andrey Nechaevskiy, Dmitry Podgainy

Copyright © 2020 for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Introduction

The use of beams of protons and nuclei of other chemical elements presents a highly promising direction for the tasks of space radiobiology (violation of the operator's functions of astronauts), as well as for the development and testing of drugs for prevention and relief of post-radiation effects. At present, LRB JINR actively carries out experiments to study the biological effects of ionizing radiation on the central nervous system of an experimental animal at different periods after irradiation [1]. In the course of the studies, a large amount of heterogeneous data (images, videos, text descriptions) from biomedical experiments appears.

The analysis of a large number of images obtained during biomedical studies is a priority objective related to computer diagnostics. The following tasks are to be solved for this:

- creation of a database for acquiring, storing and systemizing experimental data corresponding to biological objects of the study;
- elaboration of algorithms for processing experimental data on the basis of machine and deep learning methods for the tasks of pathomorphology and behavioral analysis in the study of the effects of ionizing radiation and other factors;
- development of an information system based on modern IT solutions, including web technologies, developed algorithms for statistical analysis and deep learning models, modern solutions for inference and components for the visualization of data analysis.

The transition to advanced digital technologies, such as machine learning, neural network approaches and high-performance computing technologies, for solving this class of tasks will enhance and speed up the process of analyzing and obtaining results from the data.

To process and analyze data, it is planned to use the resources of the HybriLIT [2] heterogeneous platform, which is a part of the Multifunctional Information and Computing Complex (MICC) of the Laboratory of Information Technologies of JINR [3]. The heterogeneous platform comprises the "Govorun" supercomputer [4] and the HybriLIT education and testing polygon combined by the unified software and information environment. The "Govorun" supercomputer includes two components, namely, the CPU component based on the latest Intel architectures (Intel Xeon Phi and Intel Cascade Lake processors) and the GPU component based on the NVIDIA DGX-1 Volta nodes. The NVIDIA Volta architecture is the first in the world designed specifically for machine and deep learning tasks.

The implementation of the given information system will result in the systematization of the accumulated results, the identification of hidden patterns that appear in the response of biological systems to the effects of damaging factors and, as a result, a significant simplification and acceleration of diagnosis, particularly in the early stages of diseases, of the elaboration of effective methods of prevention and counteraction to the negative effects of ionizing radiation.

2. Analysis of existing services

Medicine and radiobiology are already considered as one of the strategic and promising areas from the viewpoint of the effective implementation of machine learning, neural network approaches and high-performance computing. The application of these methods can massively increase the accuracy of diagnostics, make life easier for patients with different diseases, accelerate the development and release of new drugs, etc.

Today, platforms and online services to support decision-making in the diagnosis and analysis of different medical images exist and are developed. It is noteworthy that during the

COVID-19 pandemic, such online services have become particularly relevant and gained great popularity.

For example, HistoScan is a cloud storage of a digital archive of histological preparations, clinical information and lecture material with a high degree of data protection [5]. The given service opens up wide opportunities for remote consultations and training of specialists. HistoScan allows users to upload real-life cases from practice in their personal account, while protecting personal data. Then one can get a consultation from doctors around the world. The platform also enables remote training seminars and lectures.

The Botkin.AI online platform is a system based on artificial intelligence, it analyzes images and helps to diagnose and assess the risks of the development of diseases, including cancer [6]. For example, Botkin.AI can detect tumors on computer tomography, MRI, X-ray or mammography of a patient. It is based on its own patented technology for constructing mathematical models, with the help of which diagnostics is performed and disease risks are assessed. The program can be easily integrated with the diagnostic equipment, therefore it can be easily implemented in the work of a medical institution. Access to the system is carried out after passing the authentication procedure, after which the user can upload the necessary image in his personal account, analyze it and obtain the result based on neural network approaches.

HistoScan and Botkin.AI operate as online tools and provide the user with a web interface to interact with the software product; moreover, full access to the system is obtained only after passing authentication.

Having studied the functions and capacities of the services described above, a service architecture has been proposed, and a prototype of an online service to solve the assigned tasks has been implemented.

3. Development of the service architecture

The service is implemented as a web application and has a client-server architecture. The server is Node.js [7], and the client is written using the React.js library [8]. The service will provide convenient and safe access to the HybriLIT platform and the ML/DL ecosystem, which comprises a large number of frameworks and tools. This approach has the following advantages:

- the service provides authorized access to the computing components of the HybriLIT platform using the SLURM batch system;
- on the basis of web technologies, a common interface for researchers is implemented for interacting with the MariaDB database [9] and file storages (HLIT-storage);
- the service provides a user interface for storing, modifying and adding data obtained in the course of experiments.

In addition, all calculations will be performed on a remote dedicated server, which does not require computing resources from the user's device.

The general architecture of the service is illustrated below (Fig. 1).

The server-side (backend) is the software and hardware part of the service, which implements application programming interfaces (APIs) for interaction of the client-side (frontend) with the service components without directly accessing them from the frontend using POST requests. It simplifies work with the platform for a group of researchers and isolates different use cases depending on access rights. The core of the backend is the Node.js technology. This software platform allows one to implement a fast, asynchronous, event-driven API server. Node.js also adds the ability for JavaScript to interact with I/O devices via its programming interface and to connect other external libraries, which are written in different programming languages (C/C++, Python, etc.). The SLURM [10] scheduler

provides access to the computing power of the platform on a competitive basis. Due to these features, the API server must track the status of running tasks and keep the current state for each task in the database.

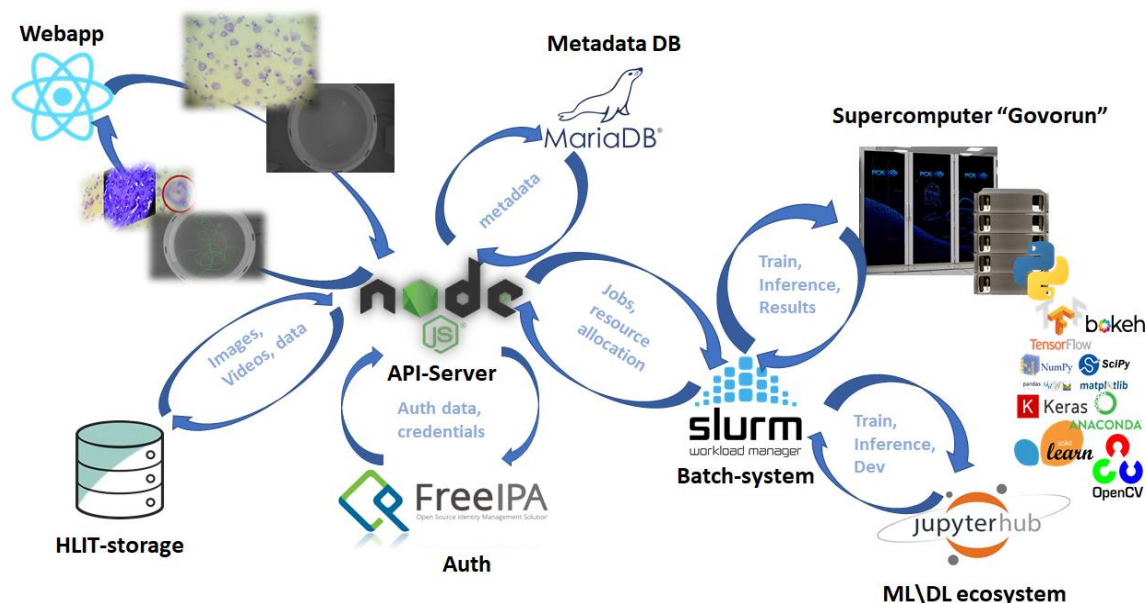


Fig. 1. General architecture of the service

To organize the storage of information about experiments, the API server under development must interact with the metadata storage implemented on the basis of MariaDB and file storages of the HybriLIT platform. Node.JS has a multitude of built-in and third-party packages to interact with these components. There is a client with non-blocking I/O MariaDB Node.js connector for interacting with the database, and there is a built-in module for working with POSIX-compatible file systems to interact with mounted file storages. An example of the simplified scheme of the API server operation is shown in Figure 2.

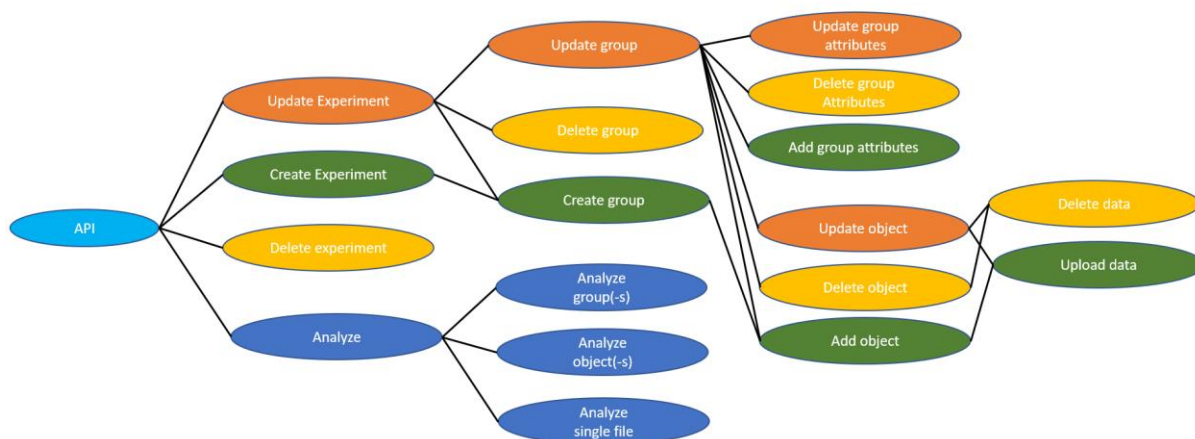


Fig. 2. Scheme of the simplified operation of the API server

The developed service presupposes a hierarchical data structure, namely, Experiment – Group – “Object” of research, which corresponds to the logic of conducting an experiment by researchers of LRB JINR.

Experiment is the major element inside of which all the contents of the experiment will be stored. It has a standard set of attributes: name, creator's name, creation date, description; and a special attribute such as exposure (for example, gamma radiation).

Group is an auxiliary element related to a specific experiment, inside of which data on "objects" of research (experimental animals) will be stored. In addition to the standard attributes, it has special ones: organ under study, dye, microscope magnification, preparation.

"Object" of research (experimental animal) is an auxiliary element related to a specific group; metadata of the "objects" of the experiment (photos and video files) will be stored inside.

The database scheme for the service under development, which is implemented in accordance with the presented hierarchical data structure, is illustrated in Figure 3. The database contains information about experiments, groups, objects, general data (creation date, creator, name, description), special attributes (organ under study, dye, exposure, microscope magnification, preparation).

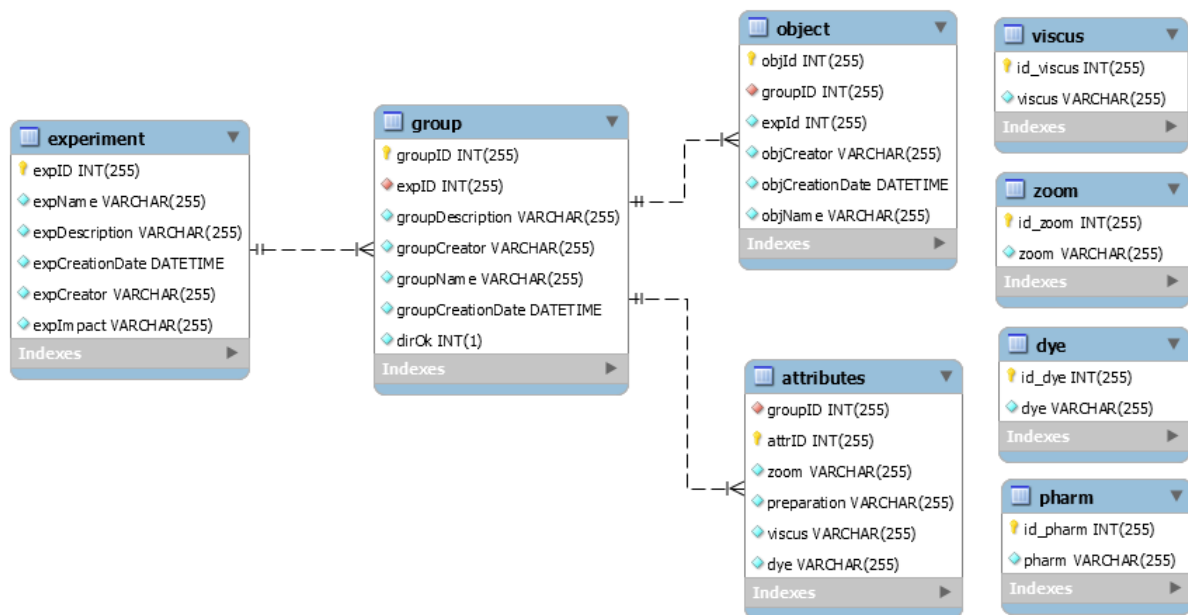


Fig. 3. Database scheme

4. Frontend and development of the service prototype

The frontend is the graphical interface that the user sees on the page. It is responsible for the appearance of the components (stylization) and their position on the page (layout). With the help of the frontend, the user can interact with the backend and the database using HTTP requests to the API. From the user's side it happens unnoticed and feels like, for example, pressing a button, entering data into a form, clicking on a page element. The following technologies have been used as means for implementing the frontend of the service:

- HTML5/SCSS;
- JavaScript, WebPack [11];
- React.js, Redux [12].

The service prototype has been implemented using these technologies. To start interacting with the service, the user needs to go through the authentication procedure by entering the login and password in the form shown in Figure 5 (*1. Authorization*).

If the login and password are correct, a valid token is issued, and the user goes to the main page of the service. The given page has a table that contains information about existing experiments in the database. Data from the database is transferred to the client as a result of an API request. In the request body, the required parameters and the token are passed to check that the user is authorized in the system, and a sample from the database in the *JSON* format appears as a response. The sample is further processed by *React.js* and displayed on the page as a table, as shown in Figure 4 (2. Page with experiments).

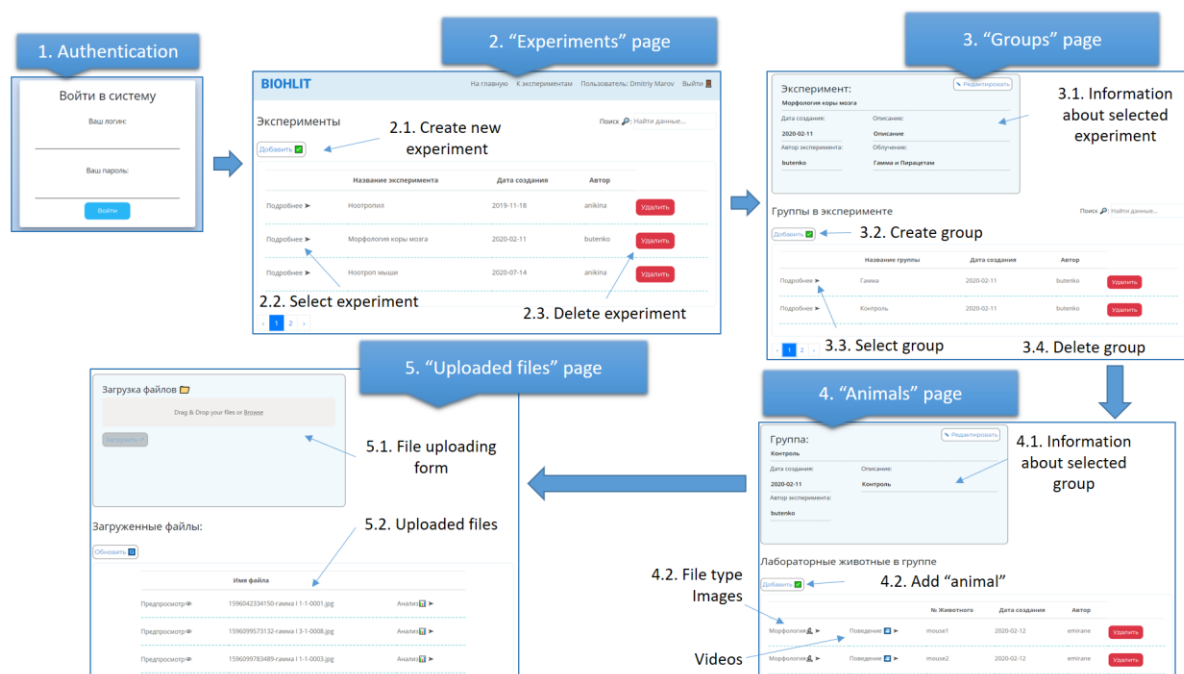


Fig. 4. Table with experimental data

The table contains basic information about experiments: name, creation date and author of the experiment. In addition, on this page, the user can manage experiments and make changes in the database, for example, add a new one or delete an existing one. To go to groups within the experiment, one should click on the "Details" button. The page with groups also provides information obtained from the database as a result of an API request. On the page with groups of the experiment (3. Page with groups in the experiment), there is a card that contains full information about the selected experiment. By clicking on the "Edit" button, the fields become active, and one can make changes, which will also be reflected in the database.

The user can create a new group and delete an existing one. To go to the page for filling in information about laboratory animals in a group, one should click on the "Details" button.

During experiments, video materials and images are accumulated in a group of pathomorphologists. To store data of a specific animal, the user must create a folder, for example, mouse1. Downloading and storing digital materials (videos and images) is also distributed on the server in the corresponding directories. For example, in the directory mouse1, two folders, namely, morphology (for images) and behavior (for videos), will be created. To download, for example, images, the user needs to click on the "Morphology" button, after that he will go to the page with a form for downloading files.

To view the image, the user needs to click on the "Preview" button in the line with the selected image. After that a window, in which the selected image is loaded from the server, opens (Fig. 5).

Название файла: http://bio.jinr.ru:8444/exp_55/grp_54/obj_36/morphology/1594713656009-control_1-2_IV-20191120105821.png

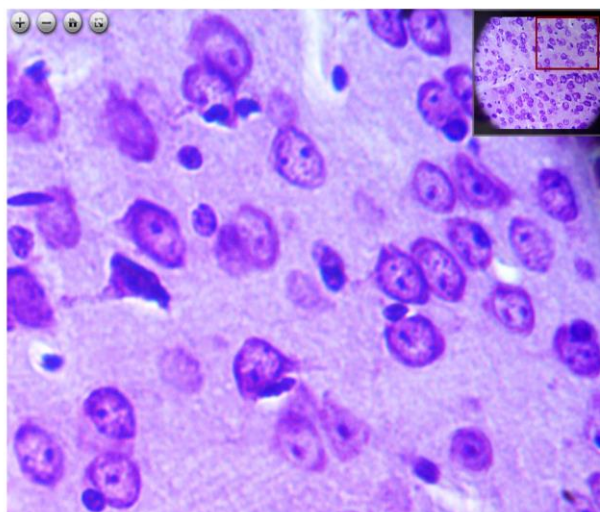


Fig. 5. Window for previewing the file

The functions of analyzing images and video materials are available to the service user. Results of image analysis can be obtained and presented in different forms. For example, one can obtain a color mask of the image, on which normochromic, lightly modified and degenerative neutrons will be marked with colors. Since the service under development is aimed not only at pathomorphologists, but also at developers, the user interface enables the view of intermediate results of algorithms, for example, primary segmentation (Fig. 6). This is necessary in order to understand how the algorithm works at each stage and to correct its work. The result of the work of the image analysis algorithm will be the statistics on objects of interest to the scientist.

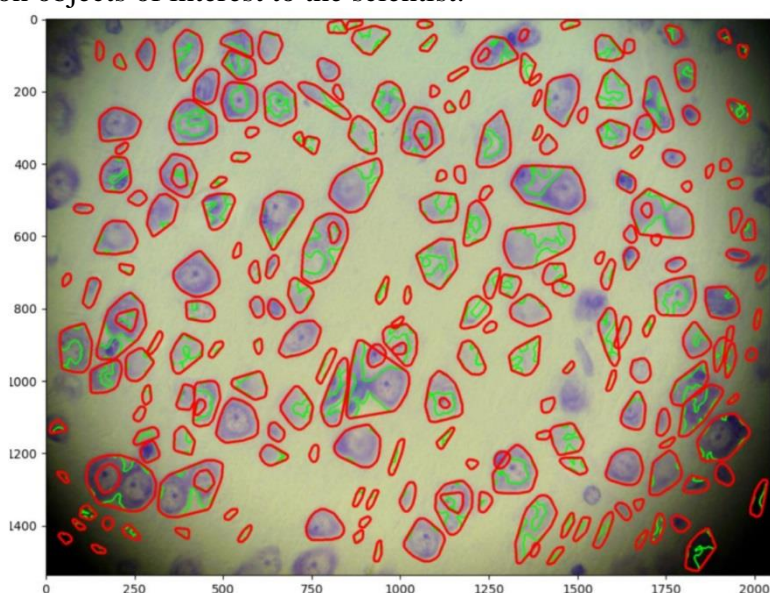


Fig. 6. Example of image segmentation

Conclusion

Thus, the article presents a prototype of a web service for conducting radiobiological studies on the HybriLIT platform. A client-server architecture has been implemented, and a database has been designed. The interaction with the service is performed using the developed user interface. Currently, the prototype provides the ability to store, create and modify data, as well as to store videos and images obtained as a result of experiments in a structured form. The present results confirm the relevance of the current project, which can be considered as a new promising solution for radiobiological studies underway at LRB JINR.

In the near future, it is planned to develop a module for presenting results of the analysis of images and video files after applying neural network algorithms, as well as to refine and enhance the user interface.

References

- [1] Kolesnikova I.A., Budennaya N.N., Severiukhin Yu.S. et al. Effect of proton irradiation on the morphofunctional state of the rat brain [Electronic resource] // http://brainres.ru/work/Sudak%202019_15%20Congress%20Proceeding_electron%20version.pdf#page=227
- [2] Adam Gh., et al. IT-ecosystem of the HybriLIT heterogeneous platform for high performance computing and training of IT-specialists // Selected Papers of the 8th International Conference «Distributed Computing and Grid-technologies in Science and Education» (GRID 2018), Dubna, Russia, September 10-14, 2018, CEUR-WS.org/Vol. 2267
- [3] MICC - INFRASTRUCTURE, n.d. Available at: <https://micc.jinr.ru/?id=29> (accessed 15.08.2020)
- [4] Supercomputer “Govorun” – Heterogeneous cluster | LIT/JINR, n.d. Available at: <http://hlit.jinr.ru/en/> (accessed 15.07.2020)
- [5] Histoscan.com is a fast and comfortable way to receive and make a second opinion for scanned digital slides. Electronic resource: http://www.biovitrum.ru/glavnaya/press-center/all_news/histoscancom/
- [6] Platform for the analysis and processing of medical images using AI technologies. Electronic resource: <https://botkin.ai/>
- [7] Node.js is a JavaScript runtime built on Chrome's V8 JavaScript engine. Electronic resource: <https://nodejs.org/>
- [8] React is a JavaScript library for building user interfaces. Electronic resource: <https://reactjs.org/>
- [8] MariaDB Server: The open source relational database. Electronic resource: <https://mariadb.org/>
- [9] Slurm Workload Manager - Overview, n.d. Available at: <https://slurm.schedmd.com/overview.html> (accessed 18.09.2020)
- [10] Webpack is a static module bundler for modern JavaScript applications. Electronic resource: <https://webpack.js.org/>
- [11] Redux is a predictable state container for JavaScript applications. Electronic resource: <https://redux.js.org/>