IMPLEMENTATION OF THE «DIGITAL LAB» PLATFORM FOR THE PROCESSING AND ANALYSIS HETEROGENEOUS NEUROBIOLOGICAL DATA IN THE NRC «KURCHATOV INSTITUTE»

A.N. Polyakov[1], I.M. Enyagina[1], D.S. Kokovin[1], M.V. Kalmikova[1]

[1] *NRC "Kurchatov Institute", 1 Akademika Kurchatova sq., Moscow, 123182, Russia*

E-mail: irina_enyagina@mail.ru

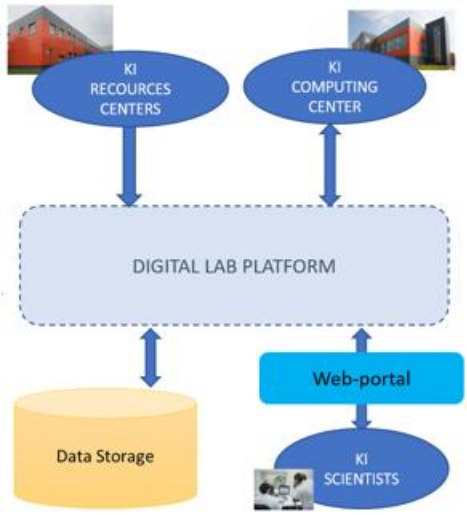
This paper presented the project of «Digital Lab» Platform implementation for organize storage, processing and analysis of heterogeneous neurobiological data (MRI, fMRI, EEG). The «Digital Lab» Platform is an information system for intensive work with metadata in heterogeneous data store with the function of the dynamic modification by user of both the data and the metadata models. There is growing interest in combining EEG, MRI and fMRI experimental data for study of human brain functional connectivity. To implement this methodology was developed the new Module of «Digital Lab» Platform for automatic processing, analysis and storage of heterogeneous neurobiological data (MRI, fMRI, EEG), obtained at NRC «Kurchatov Institute. The creation of this Module allows organizing interaction between the Resource Center «Cognimed», the Computer Center Complex for Simulation and Data Processing and the scientists at the Kurchatov Institute (Moscow). The Module uses the cluster HPC4 of the KI Supercomputer as a computing resource. The implementation of this Module allowed scientists to combine experimental EEG, MRI and fMRI data to study the functional connectivity of the human brain.

Keywords: Data analysis, Digital Lab, neuroimaging, MRI, fMRI, EEG

Andrey Polyakov, Irina Enyagina, Dmitry Kokovin, Maria Kalmikova

1. Introduction

Integration and analysis of heterogeneous (obtained at disparate equipment) experimental data is very actual question today, because it allows to get of higher-level scientific results [1], [2], [3]. Kurchatov Institute Resource Centers have more than 400 units of modern experimental devices and installations, what allows to obtain a various kind of scientific data. Kurchatov Institute Computing Center includes a Complex for Simulation and Data Processing for Mega-science Facilities providing of high-performance computing for Big Scientific Data analytics. There is a general task is to organize of better collaboration between the KI Computing Center and KI Resources Centers for improving of processing and analysis experimental scientific data. Another important task is to integration of heterogeneous (obtained on various equipment) experimental data for getting of higher level scientific results. The third task is are creating of a centralized unified data storage, which will allows use heterogeneous experimental data many times for various scientific studies. To solve these tasks we are propose apply the «Digital Lab» Platform, which allows to organize the interaction between the Resource Centers, Computing Center and scientists at the Kurchatov Institute (Figure 1).



Supercomputing clusters – 5 generations	
HPC 1 – 34 Tflops (2007)	INTEL Xeon
HPC 2 & HPC 3 – 300 Tflops (2009 & 2013)	INTEL Xeon Nvidia
HPC 4 – 500 Tflops (2015)	INTEL Xeon Haswell
HPC 5 – 600 Tflops (2015)	NVidia Tesla K80 Kepler

Figure 1. Implementation of the «Digital Lab» Platform for the processing and analysis heterogeneous neurobiological data in the NRC «Kurchatov Institute»

2. «Digital Lab» Platform

The «Digital Lab» Platform [4] is the information system for intensive work with metadata in heterogeneous data store with a function of the dynamic modification by user of both the data and the metadata models. «Digital Lab» Platform performs a role of data flows manager for organizing the conversion and transfer of data between separate elements of System for data processing and analysis. Based on the «Digital Lab» Platform it is possible to create as the simple modules for processing and analysis homogeneous data, and complex modules for processing and analysis heterogeneous data.

Key functions:

- Design of new data and metadata models with function of their dynamical modification;
- Registration of input data with extraction of initial metadata;
- Multi-criteria search on metadata attributes;
- Easy integration of external applications and services by users;
- Ingesting, restructuring and data processing jobs and workflows;
- Data share and access permissions to the toolkit resources;
- Collaborative work for multiple users.

3. «Digital Lab» Architecture

Key component is Metadata Storage, main elements of this storage are data and metadata models. Interface to these models have standardized format and will be provided by services. Services are at of three groups system, application and user's. One of the basic functions for system services are extracting initial set of metadata from input data as well creation of further metadata during the user work. System services are provide also the job loading on computing resources, processing inquires to the storage (file) system etc (Figure 2).

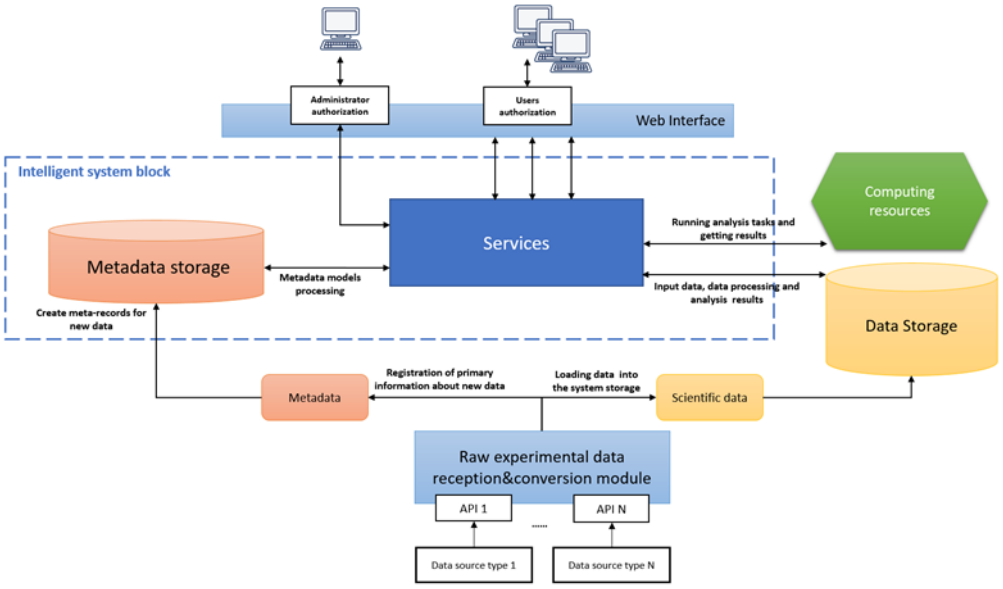


Figure 2. «Digital Lab» Platform Architecture

4. Module «Neuroimaging» for processing and analysis of heterogeneous neurobiological data

Understanding of the human brain architecture and its neuronal functional connectivity is an important neuroscience goal. There are two areas of research: anatomical structure (MRI) and functional connectivity (fMRI, EEG). Results of these researchers are used in machine intelligence and bioinformatics scientific studies (brain-computer interface, thought recognition, polygraph, etc.), and in clinical practice (neuropsychology, neurology, neuroradiology, neurosurgery, psychiatry, rehabilitation). There is growing interest in combining EEG, MRI and fMRI experimental data for study of human brain functional connectivity [1], [2], [3]. To implement this methodology the new Module «Neuroimaging» of Digital Lab Platform was developed for automatic processing, analysis and storage of heterogeneous neurobiological data (MRI, fMRI, EEG) of the Kurchatov Institute Resource Center of Nuclear Physical Research Methods "Cognimed" (Figure 3).

The Module «Neuroimaging» has the following structure:

- Data Storage: Unified storage of data and results of their processing and analysis.
- Computing Module: Data processing and analysis using specialized software (standard tools and own developed by Kurchatov Institute researchers tools, unique methods for rest-fMRI, etc.). Parallelization of calculations on supercomputer's nodes {5}.
- Analytics and Statistic Module: Data group's formation by set of parameters; analytics and statistic reports generated.
- Data Visualization Module: Creating the models of data neuroimaging with Viewer Editors.

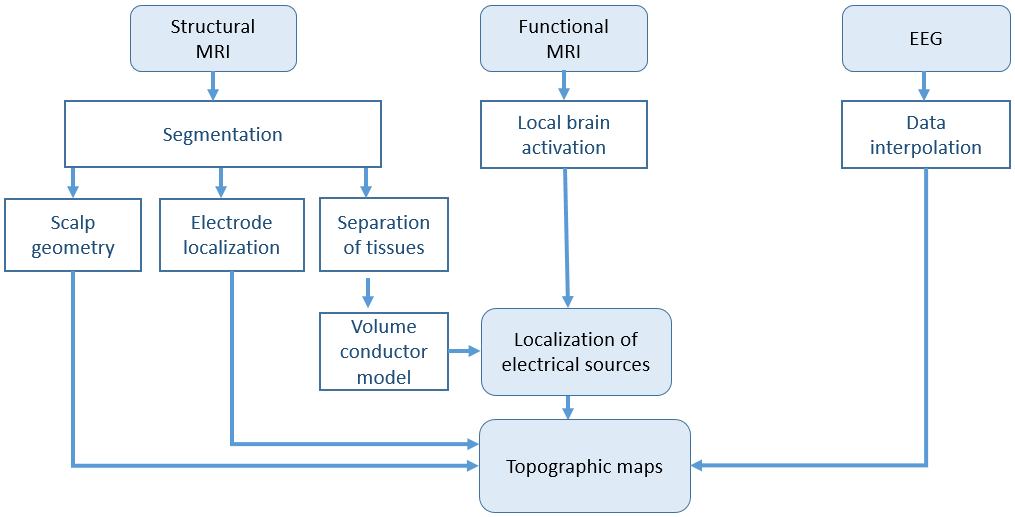


Figure 3. Automatic processing, analysis and storage of heterogeneous neurobiological data

5. Conclusion

To improving of MRI/fMRI/EEG data analysis process, we develop a Module for the processing and analysis of a MRI/fMRI/EEG experimental data, based on Kurchatov Institute «Digital Lab» Platform [4], with involvement the Kurchatov Institute Supercomputer [5]. Benefits of Module implementation: centralized data storage, with the possibility of searching on a set of parameters, what make possible in the future to reuse this data for various scientific studies; many time speed up of processing and analysis data due to parallelization of computations on supercomputer's nodes; improving of data analysis by implementing of mathematical methods developed by scientists of the Kurchatov Institute. The implementation of this Module allowed scientists to combine experimental EEG, MRI and fMRI data to study the functional connectivity of the human brain.

Acknowledgement

The work was supported by the NRC Kurchatov Institute (№1359) as research activities on the project «Creation of a distributed modular research and development platform «Digital laboratory».

References

[1] R.Abreu, A. Leal, P. Figueiredo, «EEG-Informed fMRI: A Review of Data Analysis Methods», Front. Hum. Neurosci., 2018, doi.org/10.3389/fnhum.2018.00029

[2] K.Görgenab, et al, «The same analysis approach: Practical protection against the pitfalls of novel neuroimaging analysis methods», NeuroImage, 2018, doi.org/10.1016/j.neuroimage.2017.12.083

[3] Russell A. Poldrack et al, «Handbook of Functional MRI Data Analysis», Cambridge University Press; 1 edition (August 22, 2011), ISBN-10: 0521517664, ISBN-13: 978-0521517669

[4] A.N. Polyakov et al, «Toolkit for intensive work with metadata in specialized information systems», Procedia Computer Science 119 (2017) 59–64, doi.org/10.1016/j.procs.2017.11.160

[5] V. Menon, S. Crottaz-Herbette, «Combined EEG and fMRI studies of human brain function», 2005, doi: 10.1016/S0074-7742(05)66010-2