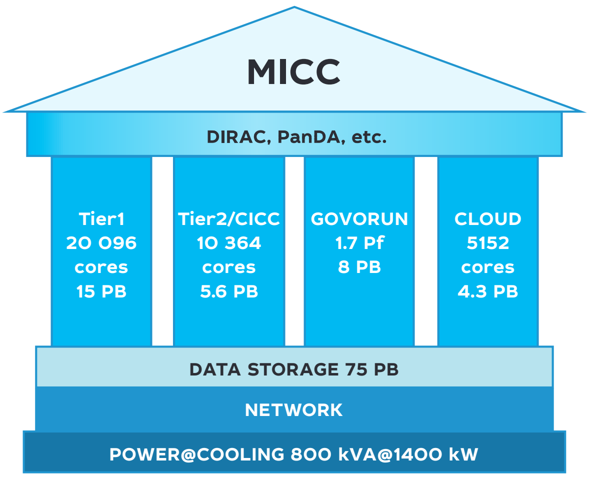
SELECTED SCIENTIFIC RESULTS OBTAINED IN 2017–2023 AT THE JOINT INSTITUTE FOR NUCLEAR RESEARCH

Dubna 2024

CONTENTS

NETWORKS, INFORMATION TECHNOLOGY AND COMPUTER PHYSICS

One of MLIT's major directions in the Seven-Year Plan for the Development of JINR for 2017–2023 was the planned development of a world-class Multifunctional Information and Computing Complex (MICC) at the Institute. It encompassed the following: the development and enhancement of the JINR telecommunication and network infrastructure; the modernization of the MICC engineering infrastructure; the modernization, development and creation of new MICC components for data storage, processing and analysis; the development of the IT infrastructure of the NICA project; the enhancement of the performance and volume of the data storage systems of the grid components, i.e., Tier1 and Tier2; the enlargement of the cloud infrastructure resources and establishment of an integrated cloud environment for the JINR Member States; the expansion of the HybriLIT heterogeneous computing complex. All indicators of the Seven-Year Plan were attained [153, 154].



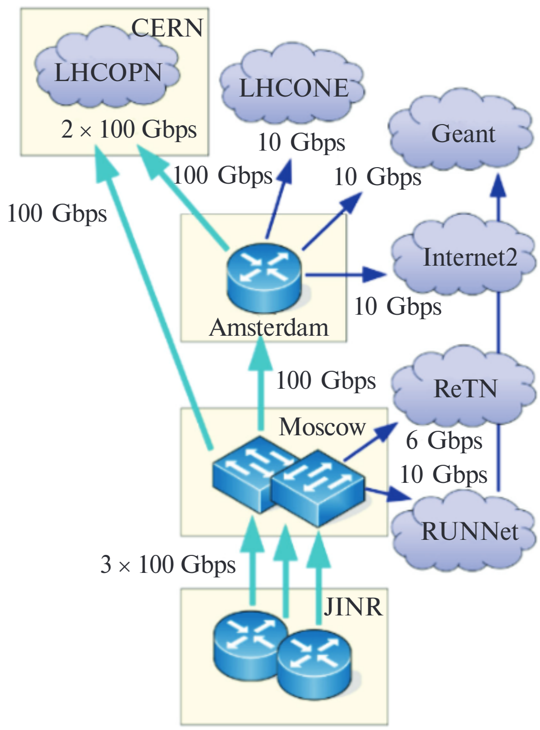
JINR Multifunctional Information and Computing Complex

JINR Network Infrastructure. The JINR network infrastructure reached a new technological level. The bandwidth of the Moscow–JINR telecommunication channel was increased from 100 Gb/s to 3×100 Gb/s, the capacity of the Institute's backbone was enhanced from 10 Gb/s to 2×100 Gb/s, and a distributed computing cluster network was built between the DLNP and VBLHEP sites with a bandwidth of up to 400 Gb/s, which meets the requirements of the NICA megaproject.

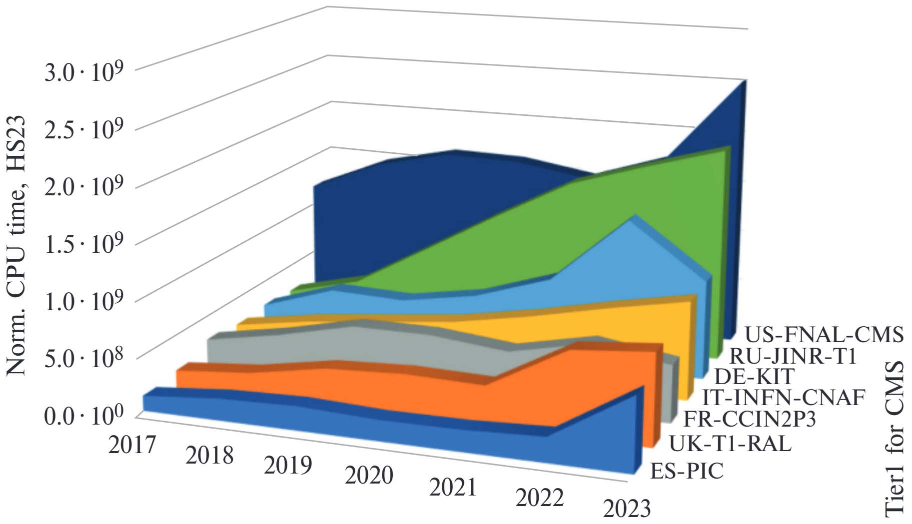
MICC Engineering Infrastructure. In accordance with the MICC project, work to modernize the morally and technically outdated engineering infrastructure of MLIT was performed. A complex of engineering systems was created, it comprises uninterruptible power supplies and two diesel generator units, representing a guaranteed power supply system that ensures the complete energy independence of the information and computing system, as well as of the network infrastructure, from the external power supply network. The modernized MICC climate control system is a complex of interconnected equipment with various air and liquid cooling schemes, with the help of which the required temperature conditions, ensuring the MICC functioning in 24×365 mode, are created [155].

JINR Grid Environment (Tier1 and Tier2 Sites). The JINR grid infrastructure is represented by the Tier1 centre for the CMS experiment at the LHC and the Tier2 centre for processing data of the NICA, LHC, BES, BIOMED, NOvA, ILC experiments, etc. Both JINR grid sites ensured 100% availability and reliability of the services.

The Tier1 processing system for CMS was expanded as planned from 3600 to 20 096 cores, cur-



JINR network infrastructure



Contribution of the world Tier1 centres to CMS experimental data processing for 2017–2023: distribution by the normalized CPU time in HS23 hours

rently delivering the 32 382.54 HS06 performance. The storage system was enlarged. The total usable capacity of disk servers was increased from 4 to 15 PB and from 5.4 to 51 PB for the IBM TS3500 and IBM TS4500 tape libraries. In terms of performance, Tier1 occupies one of the leading places among the other Tier1 centres for the CMS experiment. Since 2021, the resources of the Tier1 centre are also used to model and process data of the NICA experiments.

The computing resources of the Tier2 centre were expanded as planned from 2470 to 10 364 cores, providing the 66 788.4 HS06 performance. The total usable capacity of disk servers is 5.6 PB. The JINR Tier2 website is the best in the Russian consortium RDIG (Russian Data Intensive Grid). From 2017 to 2023, the contribution of JINR Tier2 to RDIG's productivity enhanced from 42 to 90%.

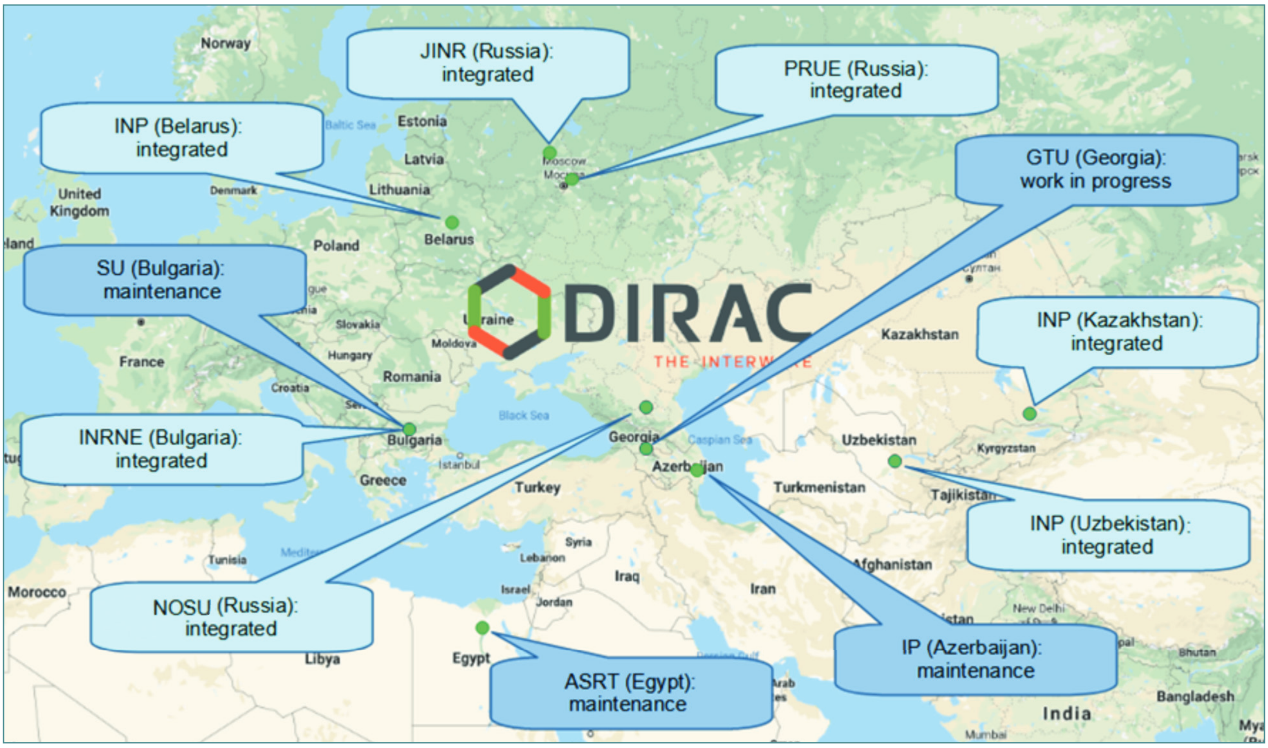
The EOS distributed storage system (the so-called "data lake") was successfully integrated into the MICC structure and is employed for storing and accessing large amounts of information. There is 23.3 PB of storage space available for EOS users. The participants of the NICA experiments, of the neutrino programme and other users store data on EOS according to quotas [156].

Cloud Infrastructure. The resources of the cloud infrastructure were enlarged from 330 to 5152 CPU cores and from 840 GB to 61.5 TB of total RAM. The total amount of disk space in the ceph-based software-defined storage was increased to 4.3 PB. The expansion of cloud infrastructure resources was financed within the NOvA/DUNE, JUNO, Baikal-GVD experiments (DLNP), the scientific groups of which are the main users of the cloud infrastructure.

During the reporting period, work to integrate the cloud structures of the JINR Member States into the DIRAC-based distributed platform was actively performed.

The cloud infrastructures of the Institute of Nuclear Physics of Kazakhstan, the Research Institute for Nuclear Problems of Belarusian State University, Plekhanov Russian University of Economics, Khetagurov North Ossetian State University, the Institute for Nuclear Research and Nuclear Energy of the Bulgarian Academy of Sciences, the Institute of Nuclear Physics of the Academy of Sciences of Uzbekistan, the Institute of Physics of the National Academy of Sciences of Azerbaijan, the Egyptian Academy of Scientific Research and Technology, Sofia University "St. Kliment Ohridski" were integrated (at present, in the last three organizations, technical work is underway, job launch has been suspended). A technical solution to create a cloud infrastructure for the Georgian Technical University is being elaborated. During the coronavirus pandemic in 2020–2022, DICE resources free from their core activity were utilized to conduct research on the SARS-CoV-2 virus within the Folding@Home platform [157, 158].

Heterogeneous Infrastructure. The Govorun supercomputer was created in 2018 on top of the experience gained during the operation of the HybriLIT heterogeneous cluster, which is part of the JINR MICC. HybriLIT has shown its relevance in solving tasks of lattice QCD, radiation biology, applied research, etc. The continuous growth in the number



Clouds of organizations integrated into the JINR distributed information and computing environment (DICE)

of users and the expansion of the range of tasks to be solved entailed not only a significant enhancement in the computing capabilities of the cluster, but also the development and implementation of novel technologies, which resulted in the creation of a new computing system, the Govorun supercomputer. The Govorun supercomputer was built as a high-performance, scalable liquid-cooled system with a hyperconverged and software-defined architecture.

The Govorun supercomputer has unique properties in terms of the flexibility of customizing the user task, ensuring the most efficient use of the computing resources of the supercomputer. The Govorun supercomputer encompasses a GPU component, a CPU component and a hierarchical data process-

GPU accelerator	Hyperconverged CPU and distributed storage nodes

Govorun supercomputer

ing and storage system with a read/write speed of 300 Gb/s, which is an extremely convenient tool for processing large data arrays, including for the NICA megaproject. According to the speed of accessing data, the storage system is divided into layers, namely, very hot data, the most demanded data, to which it is currently required to provide the fastest access, hot data and warm data. Each layer of the developed system can be used both independently and as part of data processing workflows. For the high-speed data processing and storage system, the Govorun supercomputer received the prestigious Russian DC Awards 2020 in "the Best IT Solution for Data Centres" nomination.

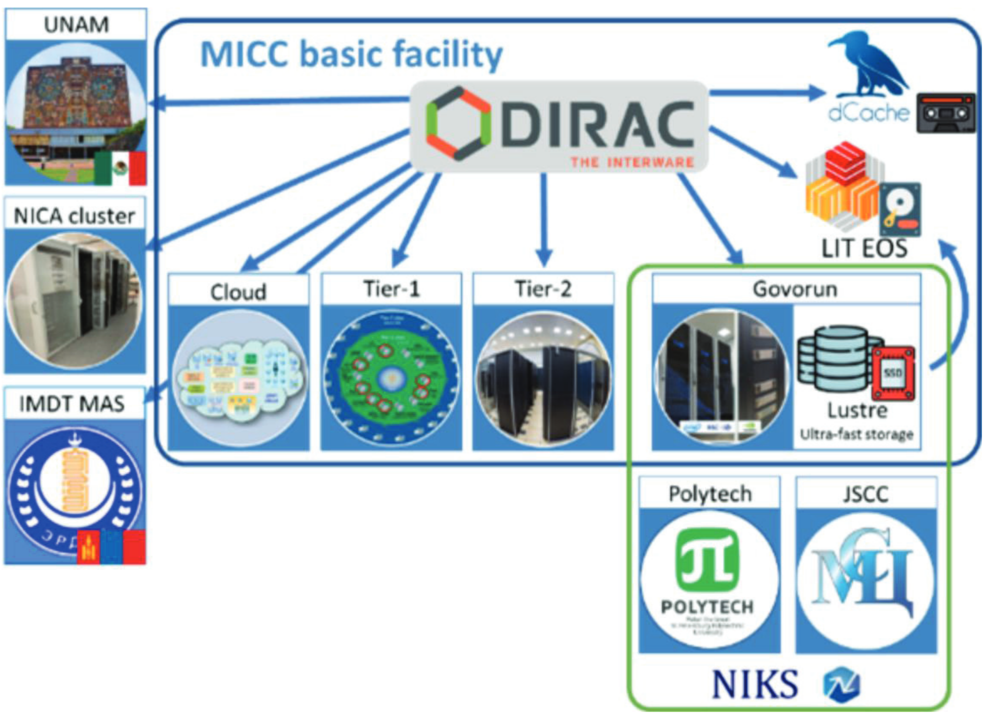
Since its presentation, the overall performance of the Govorun supercomputer has enhanced from 0.5 to 1.7 PFlops for double precision operations, and the total capacity of the hierarchical storage has increased from 288 TB to 8.6 PB.

In 2020, the development and implementation of an ecosystem for machine/deep learning and high-performance computing (ML/DL/HPC ecosystem) in the HybriLIT platform were completed. The ecosystem is actively used to create algorithms on top of neural network approaches for solving applied tasks.

Since 2021, an information and computing system (ICS) to solve tasks related to the computations of electronic shells of superheavy elements has been intensively developed on the HybriLIT platform. The ICS embraces the computing resources of the Govorun supercomputer and a set of IT solutions and software required for modeling electronic shells. To calculate the electronic properties of superheavy elements, intensive computing using the AMS and DIRAC software was performed on this system. In addition, to develop quantum algorithms, a testing polygon for quantum computing with installed quantum computing simulators, namely, Cirq, Qiskit, PennyLane, capable of operating on various computing architectures, was deployed in the ICS.

At the end of 2021, a scalable research infrastructure of a new level was created on the basis of combining the supercomputers of JINR, JSCC RAS and SPbPU. It allows the participants to enlarge their local computing power, to provide access to the means for storing and processing large data volumes, to distributed data storages (data hubs), as well as to utilize each other's capacities in the case of peak loads. Such an infrastructure is also in demand for the tasks of the NICA megascience project.

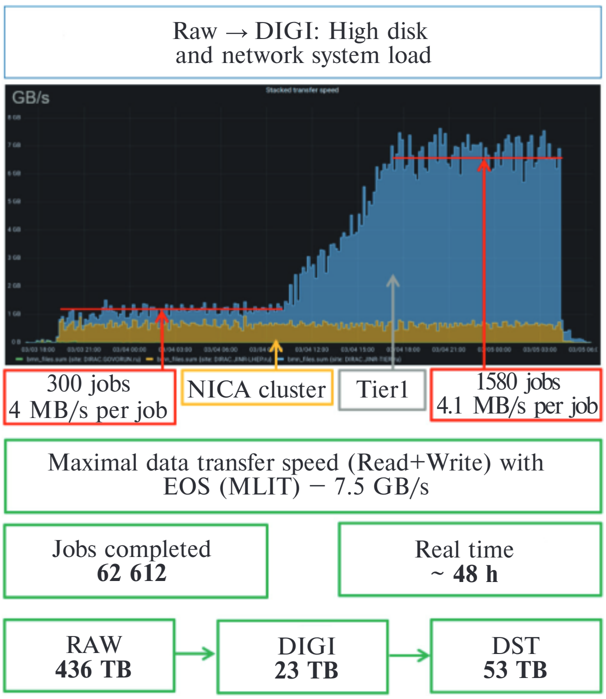
The resources of the Govorun supercomputer are employed by scientific groups from all Laboratories of the Institute. The total number of Govorun supercomputer users is currently 312, of which 255 are JINR staff members, and 57 are from the Member States. Access to the resources of the Govorun su-



Scheme for the integration of geographically distributed heterogeneous resources based on the DIRAC Interware percomputer is provided only to those users who are directly involved in the implementation of the JINR Topical Plan. The results obtained using the resources of the HybriLIT platform, including the Govorun supercomputer, the education and testing polygon and the ML/DL/HPC ecosystem [159–162], are reflected in over 300 user papers, two of them are in the Nature Physics journal.

Integration of Computing Resources. Within the implementation of the Seven-Year Plan, using the DIRAC (Distributed Infrastructure with Remote Agent Control) Interware, the computing resources of Tier1/Tier2, the Govorun supercomputer, the cloud environments of JINR and its Member States' organizations, the NICA cluster, the cluster of the National Autonomous University of Mexico (UNAM), the cluster of the Institute of Mathematics and Digital Technology of the Mongolian Academy of Sciences (IMDT MAS), the National Research Computer Network of Russia and storage resources, namely, dCache, EOS and the Lustre ultrafast data storage system, were combined.

The main users of the distributed platform are the scientific groups of the MPD, SPD, BM@N, Baikal-GVD experiments. The 8th BM@N physics run was the first time at JINR when the entire computing infrastructure, integrated by the DIRAC platform, was used for the complete reconstruction of raw experimental data [163, 164].



Data processing intensity within the 8th run of the BM@N experiment using DIRAC

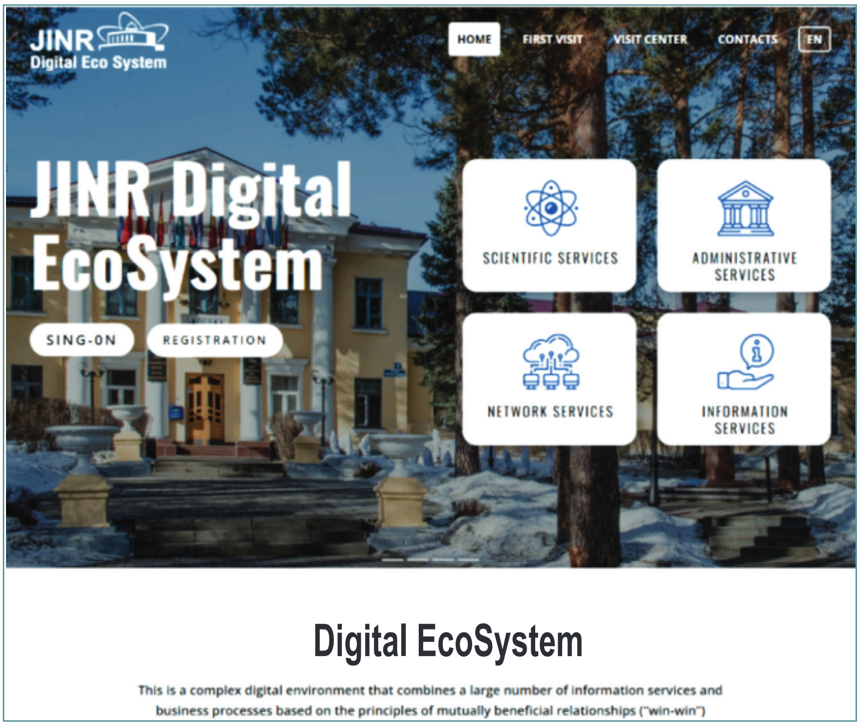
Monitoring and Accounting System. To ensure the reliable functioning of the MICC, a multilevel monitoring and accounting system, which operates in 24×365 mode, was created and is expanding.

The system performs more than 16 000 checks for over 1800 elements involved in the monitoring. The developed complex monitoring system of the MICC enables to receive information from different components of the computing complex: the engineering infrastructure, the network, compute nodes, job launch systems, data storage elements, grid services, which guarantees a high level of reliability of the MICC [165].

MICC monitoring system

JINR Digital EcoSystem. Since 2022, work to create the Digital JINR platform, i.e., the JINR Digital EcoSystem, has been underway. Its major objective is to provide a single environment for the creation and development of digital services, their integration with each other and the analysis of information on all aspects of JINR's activity. The Digital EcoSystem encompasses a wide range of services, from resources for users of basic facilities to arranging business trips, vouchers, ordering certificates, etc. The main groups of services are administrative (area of responsibility of the JINR Development of Digital Services Department) and scientific. Access to the system is based on the JINR Single Sign-On (SSO) authentication service via a single access point of the Digital EcoSystem [166].

Methods, Algorithms and Software for Modeling Physical Systems, Mathematical Processing and Analysis of Experimental Data. One of the important activities of MLIT within the implementation of the Seven-Year Plan was to provide mathematical, al-



Single access point interface of the Digital EcoSystem

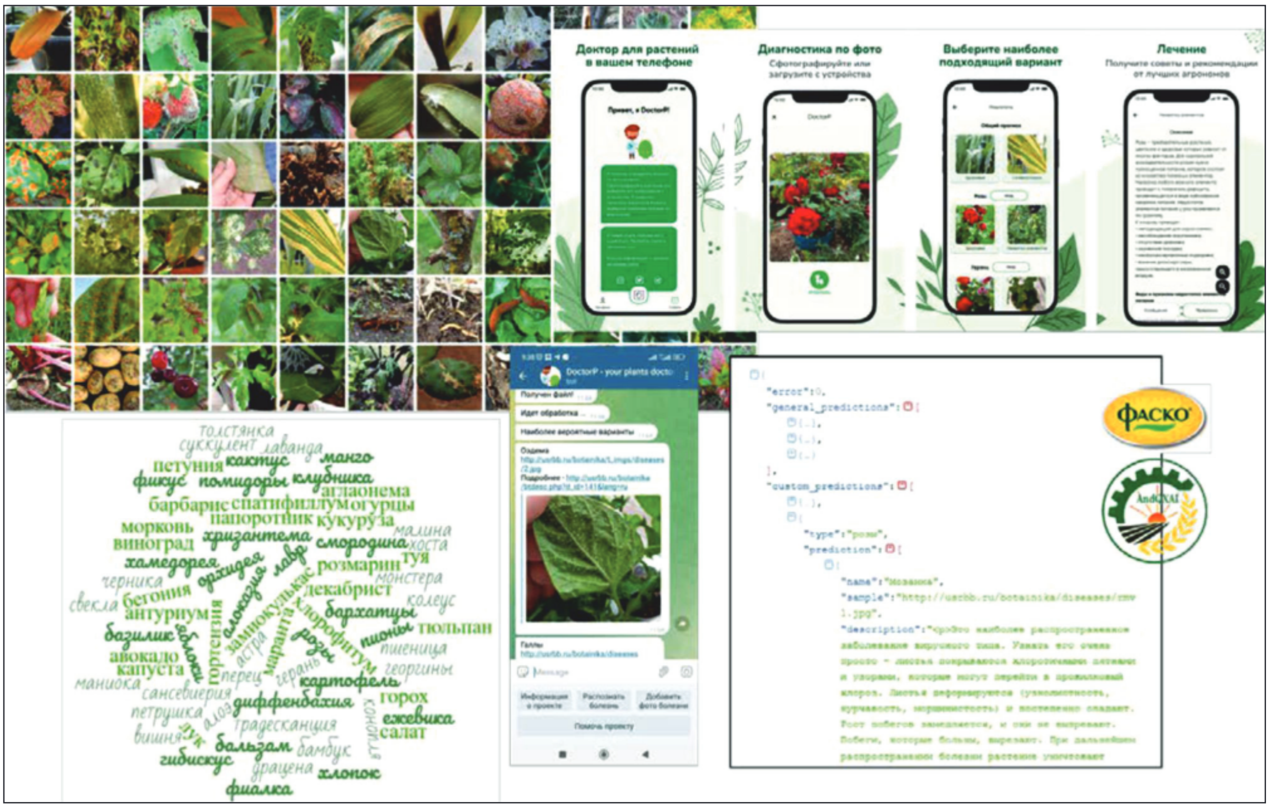
gorithmic and software support for experimental and theoretical studies underway at JINR, in many of which computing resources are an essential tool for producing significant scientific results. A summary of some prominent results is presented below.

For the first time, it became possible to algorithmize the classical problem posed by the outstanding Norwegian mathematician Sophus Lie in 1883 and verify by point transformations the linearizability of nonlinear ordinary differential equations resolved with respect to the highest order derivative. The developed algorithm enables not only to determine whether a given equation is linearizable, but also to obtain a system of differential equations for the linearizing transformation, the solution of which gives an explicit form of such a transformation. At the International Symposium on Symbolic and Algebraic Computation (ISSAC 2017, Kaiserslautern, Germany, July 25–28, 2017), the work was recognized as the best presented and was awarded a prize from the American Association for Computing Machinery.

In April 2020, the scientific paper of an international research group, in which staff members of MLIT (O. Chuluunbaatar) and BLTP (Yu. Popov) participated within JINR international cooperation, was published in Nature Physics. A kinematically complete experimental measurement of the characteristics of Compton scattering at free atoms, using the highly efficient method of COLd Target Recoil Ion Momentum Spectroscopy (COLTRIMS), was conducted. A theoretical description of the phenomenon is based on the calculations carried out on the Govorun supercomputer.

The KANTBP 3.1 program for calculating the energy values, reflection and transmission matrices, and the corresponding wave functions in the adiabatic coupled-channel approach was developed and published in the CPC Program Library. The advantage of this program compared to the widely used CCFULL program for calculating the cross section for heavy-ion fusion and fission reactions is the careful processing of the boundary conditions for solving the system of coupled Schrödinger equations, which allows maintaining a high accuracy of calculations that take into account a large number of coupled channels. The theoretical cross sections obtained with the help of the KANTBP 3.1 program well describe experimental data for different heavy-ion fusion and fission reactions.

In the course of research conducted jointly with FLNP within the UNECE International Cooperative Program (ICP) Vegetation for monitoring and predicting air pollution processes in Europe and Asia, a cloud platform for managing monitoring data was developed at JINR. To ensure the reliable storage, analysis, processing and collective use of monitoring data, modern methods of software management, statistics and machine learning were applied, which also enabled to use satellite image data



Examples of DoctorP platform interfaces

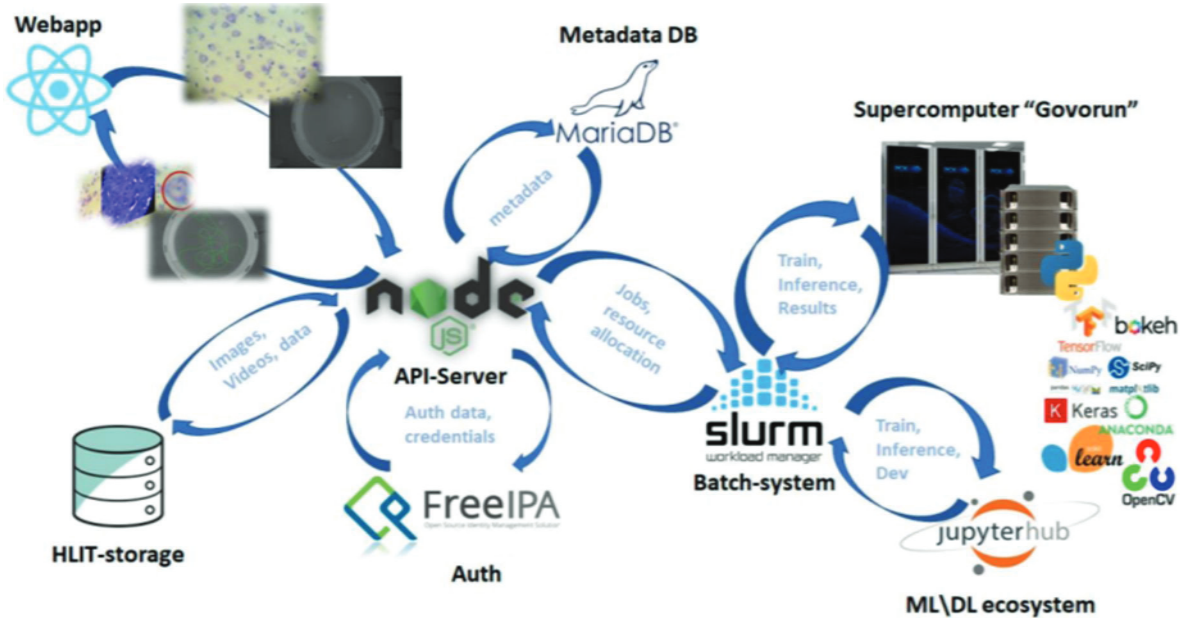
to predict atmospheric pollution by certain heavy metals in a number of European regions.

A platform and a mobile application (DoctorP) for detecting plant diseases and pests were elaborated. Both a general model capable of detecting 68 disease classes and specialized models for 30 ornamental and agricultural crops are available. The database contains over 6000 images. To obtain a prediction and treatment recommendations from experienced agronomists, one just needs to send a photo showing the problem. The platform can be accessed by third-party applications and services. Garden Retail Service (formerly Fasko) and the Andijan Institute of Agriculture and Agrotechnology (Uzbekistan) already took advantage of this opportunity.

A hardware and software platform was developed on top of quantum fuzzy controllers built into the control circuit to solve the task of controlling the pressure and flow of liquid nitrogen of superconducting magnets of the cryogenic system of the NICA accelerator complex. The multilevel control system embraces the existing lower executive level, based on the Tango Controls system, and a new level at which control actions are formed using a quantum fuzzy controller. At the same time, optimal control quality parameters, such as temperature, nitrogen consumption, speed, the required pressure level and minimal complexity of the control implementation, are provided. The operability and effectiveness of the developed intelligent system for the remote control of the technological process of cooling a superconducting magnet with a guaranteed achievement of a stable superconductivity zone were experimentally demonstrated. The design of quantum fuzzy controllers is based on quantum information technologies and is performed using the QSCIT (Quantum Soft Computational Intelligence Toolkit) software developed by MLIT specialists.

Since 2020, on the basis of the ML/DL/HPC ecosystem, a joint project between MLIT and LRB on the creation of an information system (IS) for analyzing behavioral and pathomorphological changes in the central nervous system when studying the effects of ionizing radiation and other factors on biological objects has been developing. Algorithms for experimental data processing on top of machine and deep learning methods and computer vision are implemented in the developed system. The IS embraces reliable modern tools of authentication and hierarchical data access control, a data storage system, as well as components for convenient work and the visualization of data analysis results.

Within the ML/DL/HPC ecosystem, using the example of solving a specific problem of investigating the magnetization dynamics in a Josephson ϕ^0 junction, a methodology for developing software



Architecture of the information system for radiation biology tasks

modules based on JupyterHub, which enable not only to perform computations, but also to visualize the results of the investigation and to accompany them with required formulas and explanations, was presented. A parallel implementation of the algorithm for carrying out computing for various values of model parameters on top of the Joblib Python library as well as modules with integration of Matlab code into Jupyter Notebook, which enable efficient applied computations for image analysis, were developed.

In 2017–2023, more than 1300 scientific papers within research conducted by MLIT staff members and over 700 articles within international collaborations were published, about 800 reports were delivered at international and Russian conferences. All studies conducted correspond to the expected results stated in the Plan [167–174].