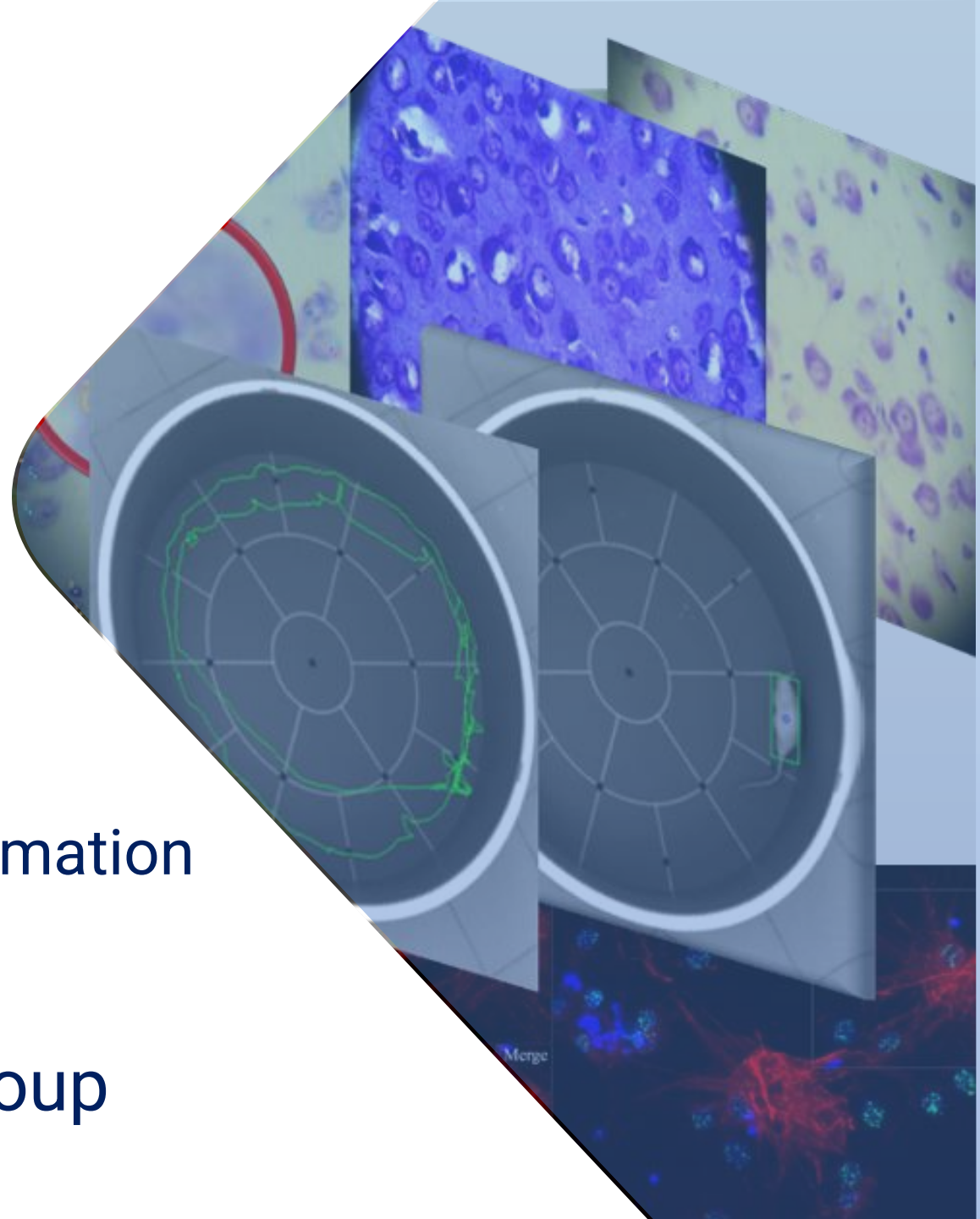# IT ecosystem based on machine learning methods and data analysis technologies for radiobiological research

O.I. Streltsova
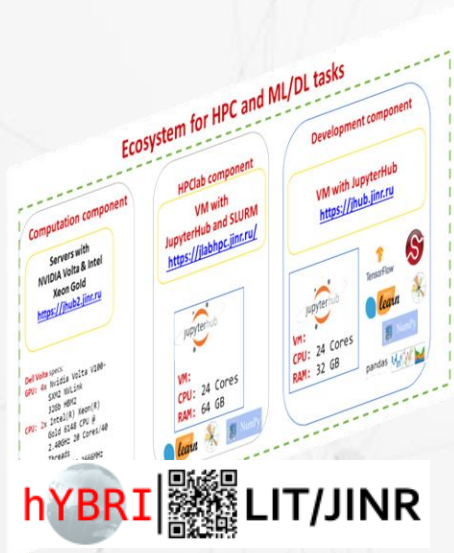
Meshcheryakov Laboratory of Information Technologies, JINR

On behalf of the **BIOHLIT** group
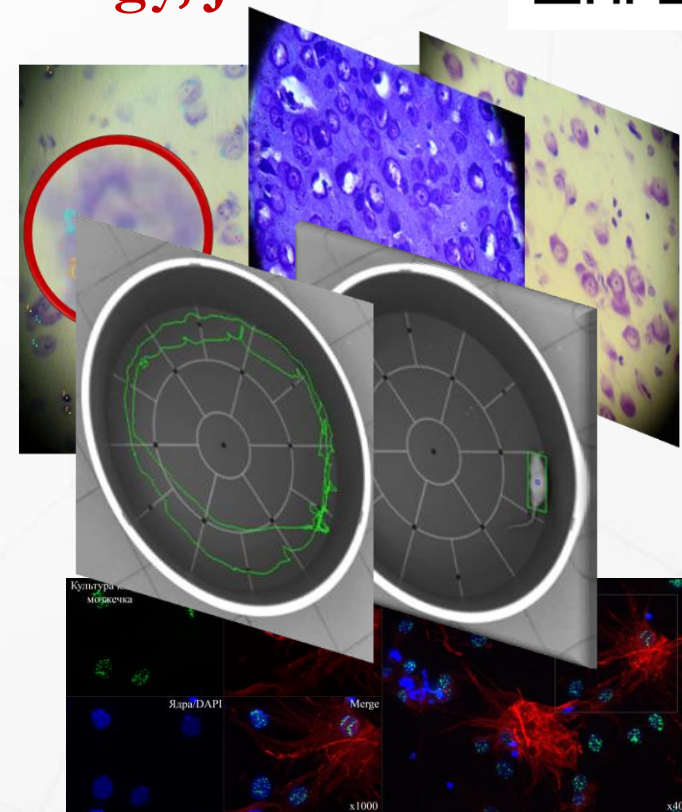
BIOHLIT group

Meshcheryakov Laboratory of Information Technologies, JINR

Anikina A.I.
Butenko Yu.A.
Zuev M.I.
Nechaevskiy A.V.
Podgainy D.V.
Stadnik A.V.
Streltsova O.I.

Laboratory of Radiation Biology, JINR

Boreyko A.V.
Zadnepryanec M.G.
Kolesnikova I.A.
Lalkovicova M.G.
Lyakhova K.N.
Severiukhin Yu.S.
Utina D.M.
Chausov V.N.
Chramko T.S.

In collaboration with
Khetagurov
North Ossetian State University (NOSU)
Coordinator: Tvauri I.V.

In collaboration with
University of Belgrade
Coordinator: Dr Marko Ćosić

In collaboration with
Germany:
Streltsov Alexei I.
Gromov E.

Complex analysis of the obtained results at all stages of the study
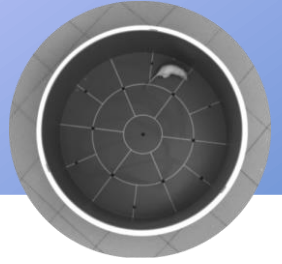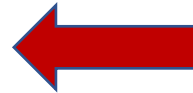
Stage #1

Stage #2
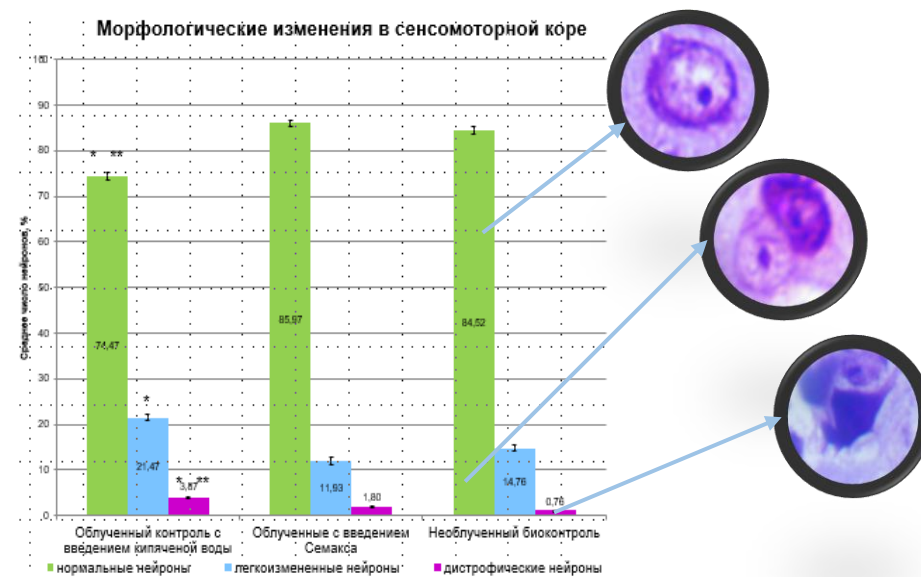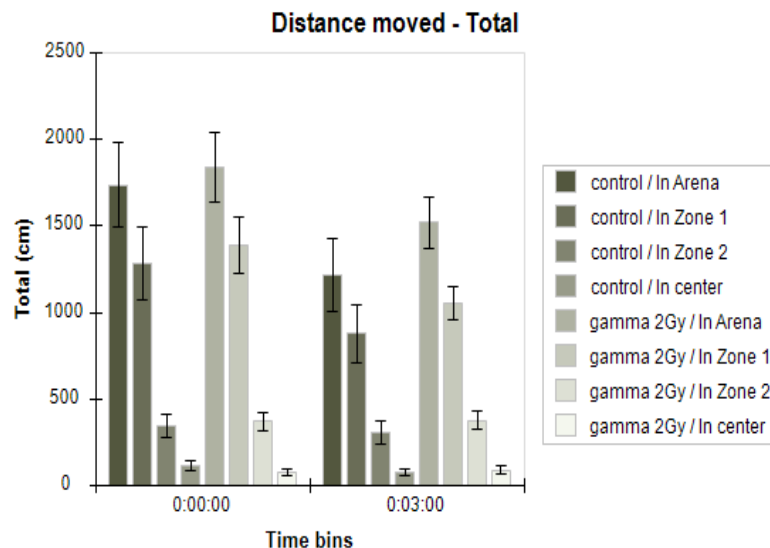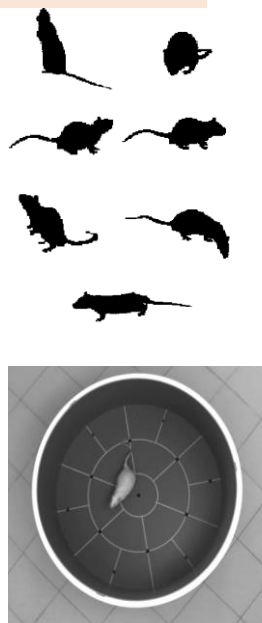
The IS is necessary for:

- **Storing** the entire array of experimental data (photo and video materials, pdf, excel, doc files with information about the experiment) and minimizing the risk of information loss.
- Convenient access for all members of the research group and ability to conduct both step-by-step and comprehensive data analysis in a unified information space.
- **Reducing** the time spent on data processing, increasing the speed of obtaining qualitative results and reducing the subjectivity of the experimental data processing approach.

BIOHLIT

The first stage is related to the analysis of behavioral reactions of laboratory animals exposed to ionizing radiation, as well as of the possible effect on the behavior of pharmacological preparations.

From the point of view of data analysis, this stage is connected with the collection and analysis of **video data** reflecting behavioral reactions of laboratory animals during experimental tests using specialized stands.

Behavioral laboratory room (LRB JINR)
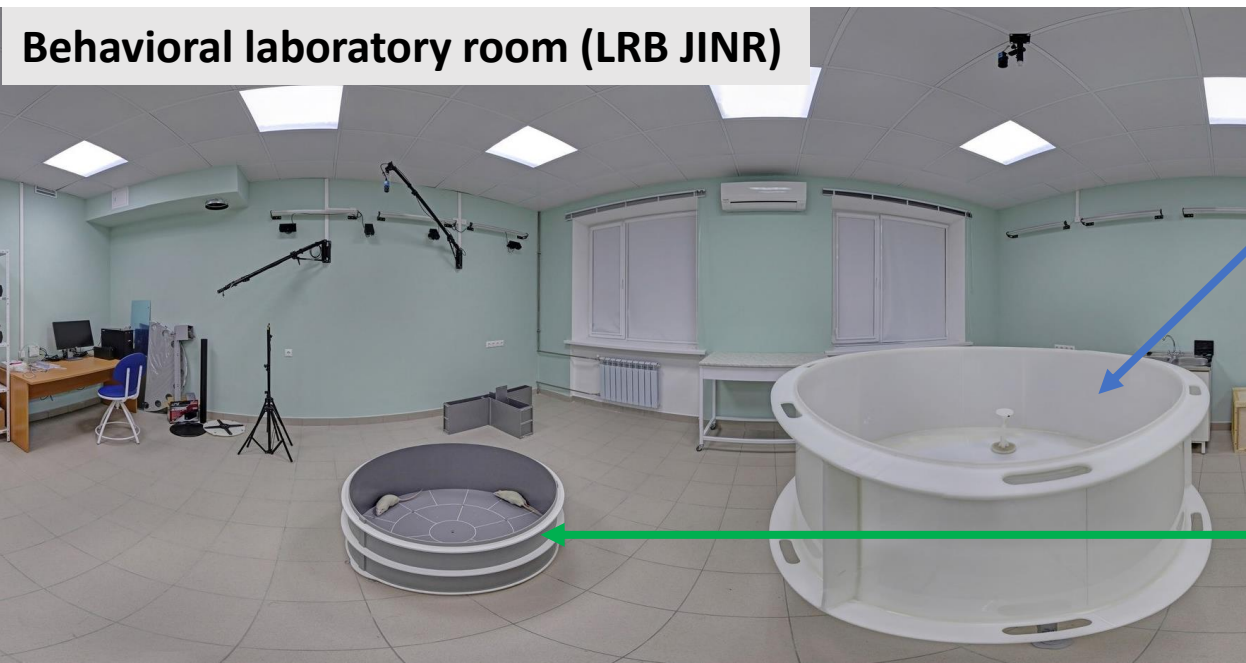
Test system "Water Maze"

Test system "T-maze"

Test system "Open Field"

"Beam-walking" (SPC Open Science, Russia)

Test system "Open Field"

ML/DL/computer vision algorithms

The **second stage** is aimed at studying morphological changes in the cells of the central nervous system to establish a connection with the behavioral reactions of laboratory animals obtained during the first stage of investigations.
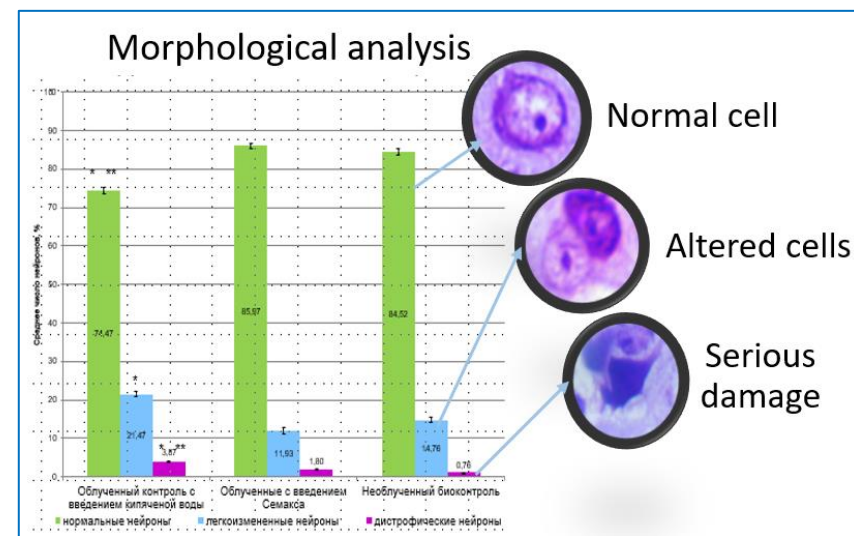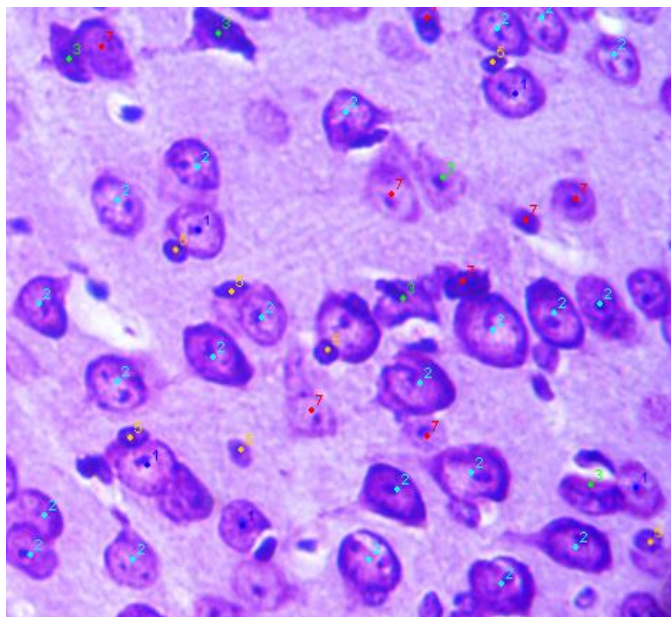The histological examination of other organs of laboratory animals is also carried out.
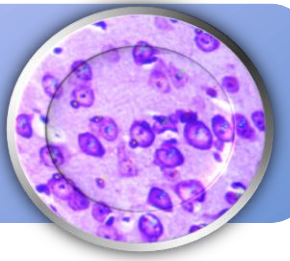
From the point of view of data analysis, this stage is connected with the analysis of **images of slices** of the nervous tissue of different parts of the brain obtained using a light microscope, a camera and programs that allow counting cells with certain changes.

The problem of automating the morphological analysis of histological preparations is solved within the project by implementing algorithms based on the neural network approach and computer vision methods.

Morphological analysis

Normal cell

Altered cells

Serious damage

Облученный контроль с введением кипяченой воды
Облученные с введением Семакса
Необлученный биоконтроль

нормальные нейроны   легкоизмененные нейроны   дистрофические нейроны

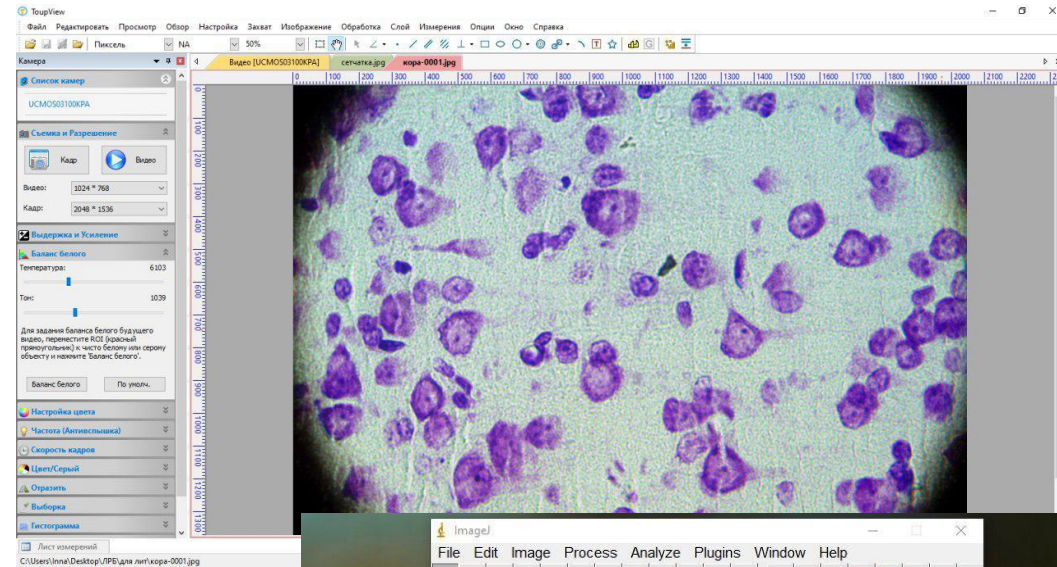# Stage #2: studying morphological changes in the cells of the central nervous system

Stages of histological research

Light microscope

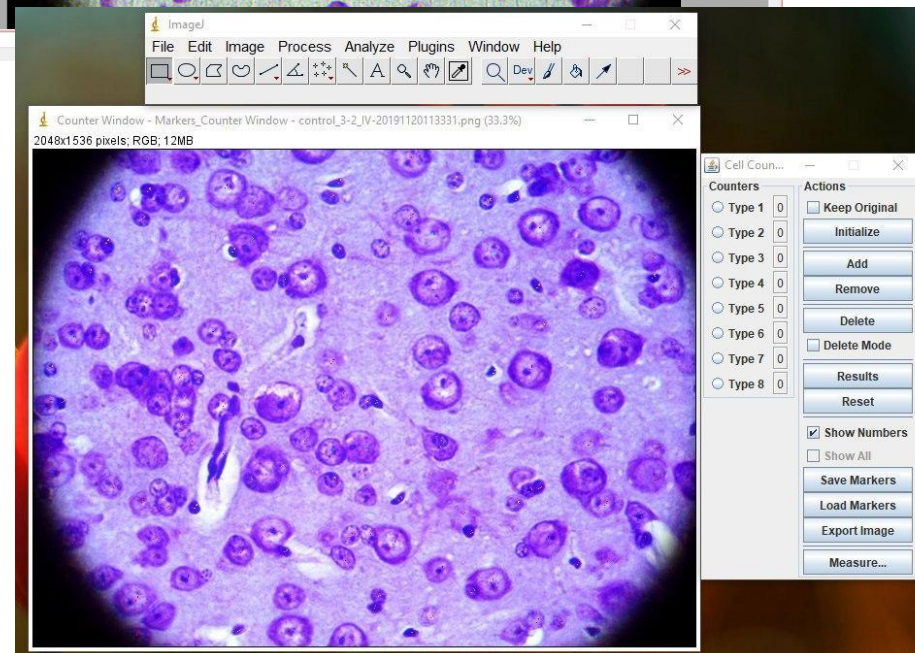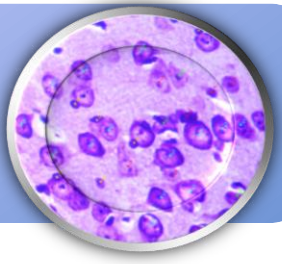**M. 40x10. Histological staining of the nervous tissue by the Nissl method**

ToupView

Histological preparations

ImageJ

Used classification of brain cells:

❏ normal (1);

❏ easily modified (2): morphofunctional and compensatory-adaptive;

❏ degenerative (3);

❏ glia (5);

❏ elements of the nervous tissue that are difficult to identify (7)
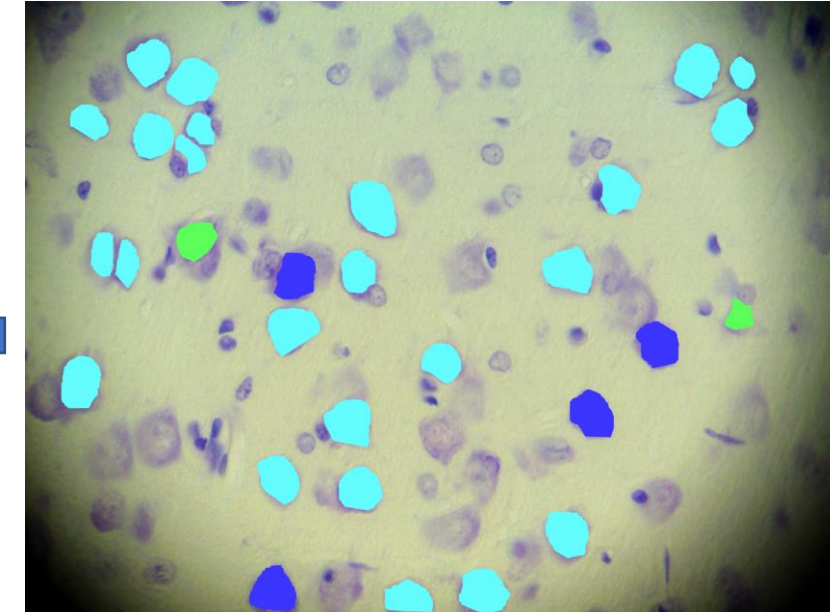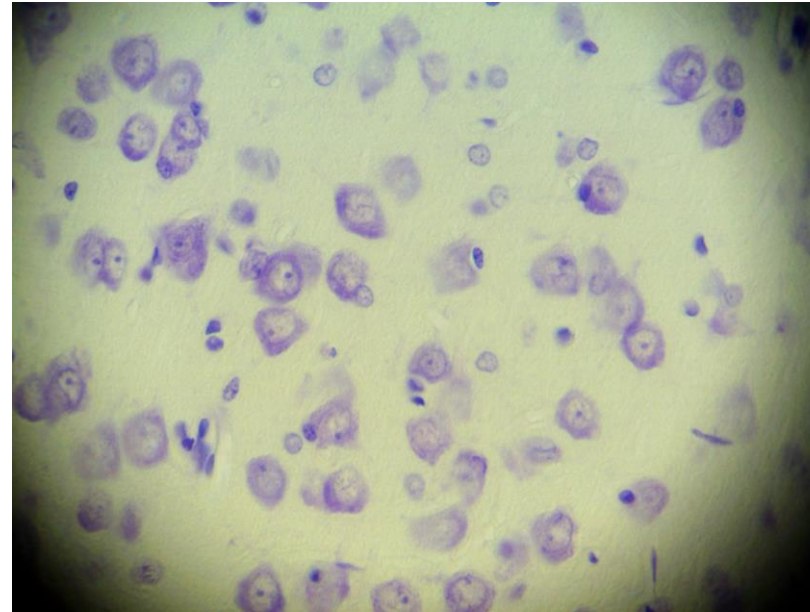
For the segmentation task, the U-net neural network architecture is well suited

Stages of histological research

Automation options:

1. automation of the morphological analysis of brain cells of experimental animals using photos of micropreparations
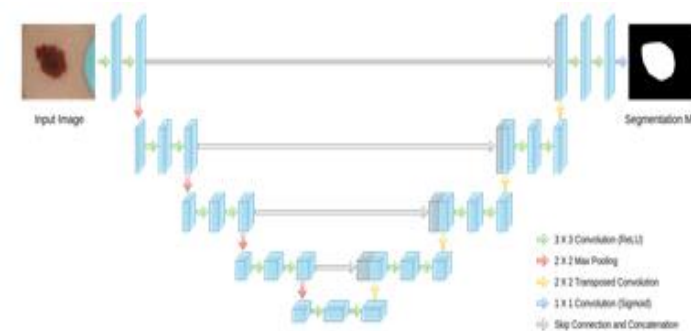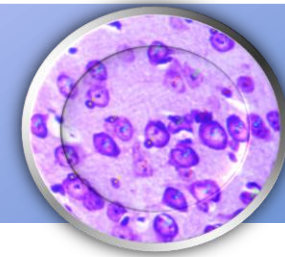2. storage of photo images before and after their analysis by an expert

Morphological analysis

Normal cell

Altered cells

Serious damage
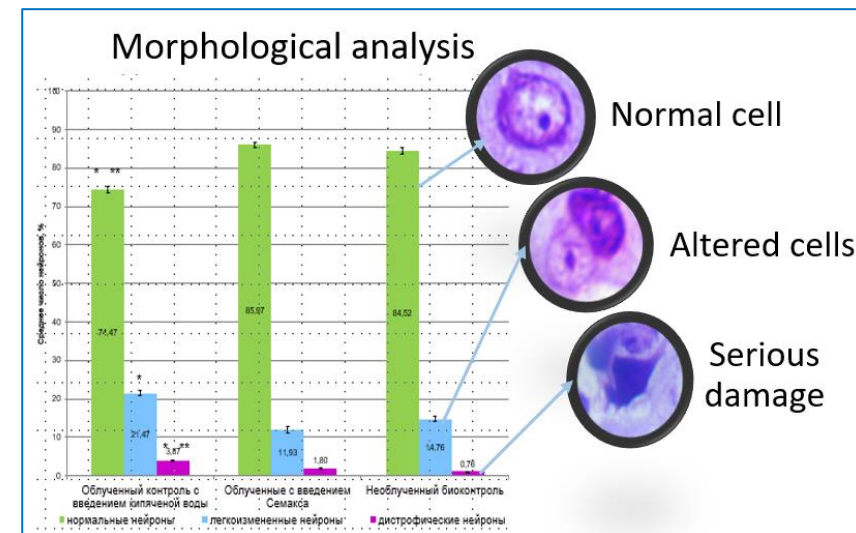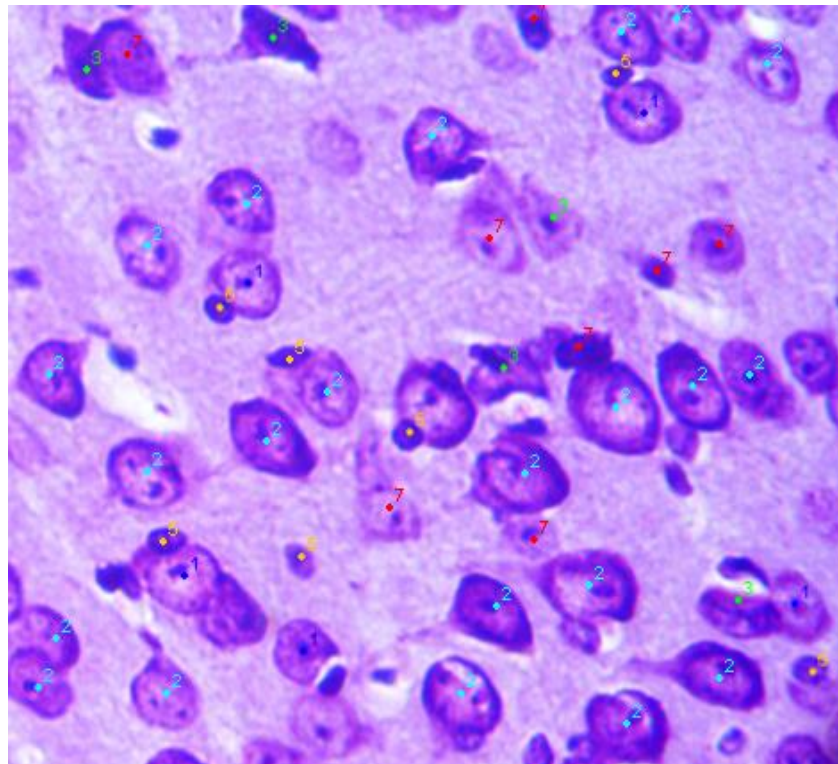
For the segmentation task, the U-net neural network architecture is well suited
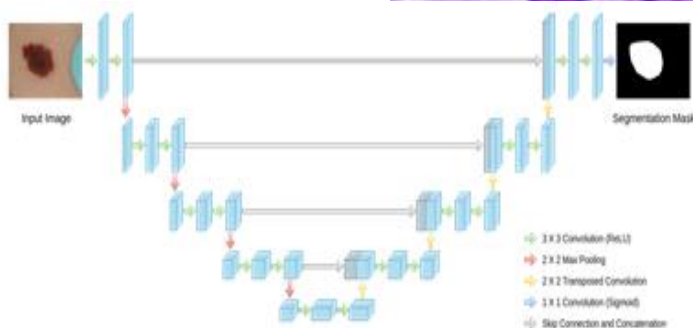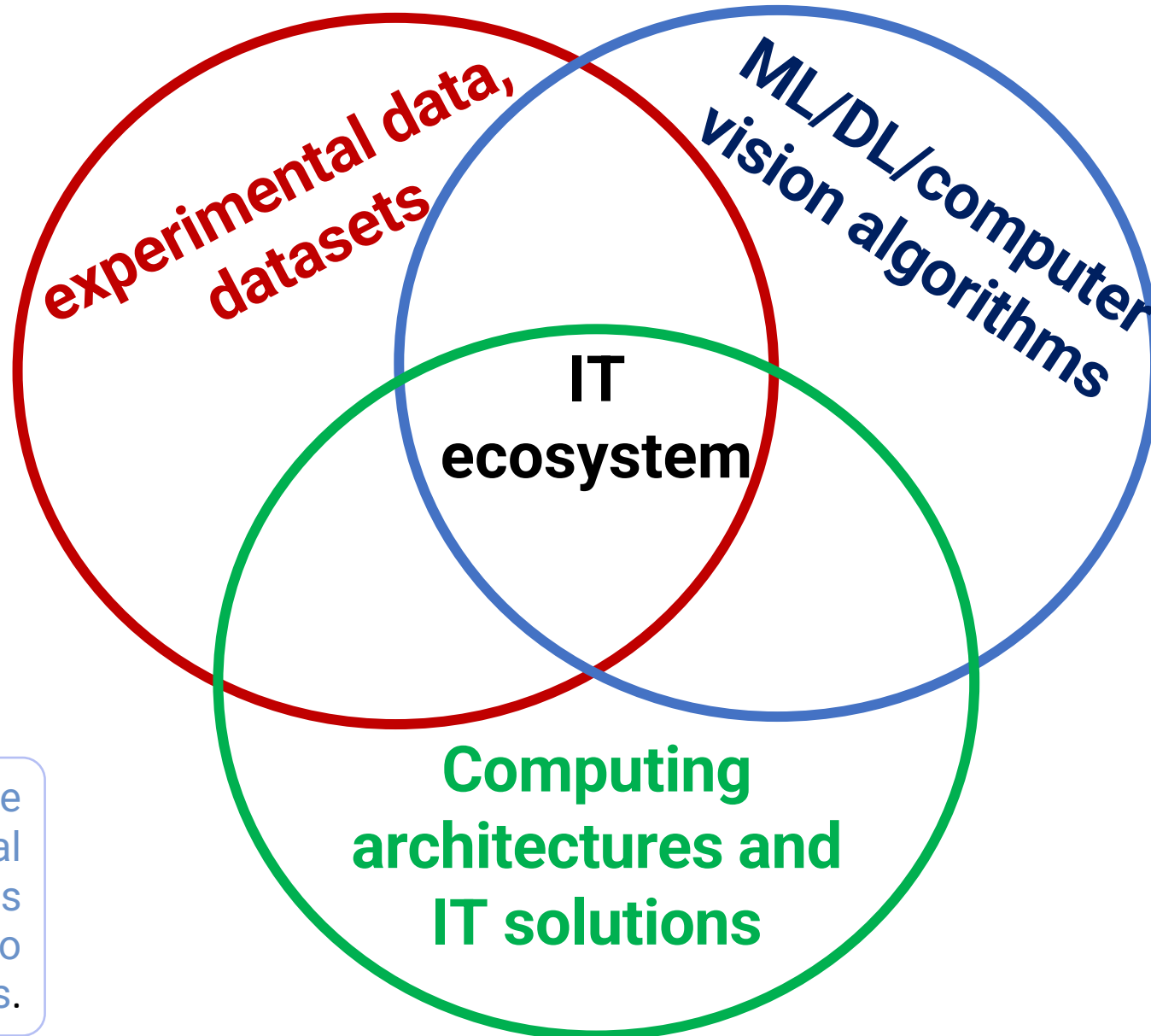
Venn diagram on the implementation of the neural network approach, methods and algorithms of ML/DL to solve applied tasks.

IT ecosystem: computing architectures and IT solutions

MICC component: HybriLIT platform

Unified software-hardware environment

VM pool – User interfaces

VM-UI (ssh) | VM-UI (ssh) | VM-UI (HLIT-VDI) | VM-UI (HLIT-VDI)

Data storage system

Computational field

Supercomputer "Govorun" | Education and testing polygon "HybriLIT"

VM pool – for management and information services

Website | GitLab

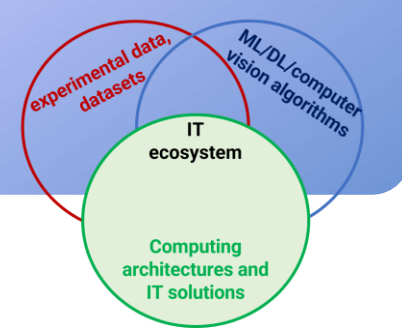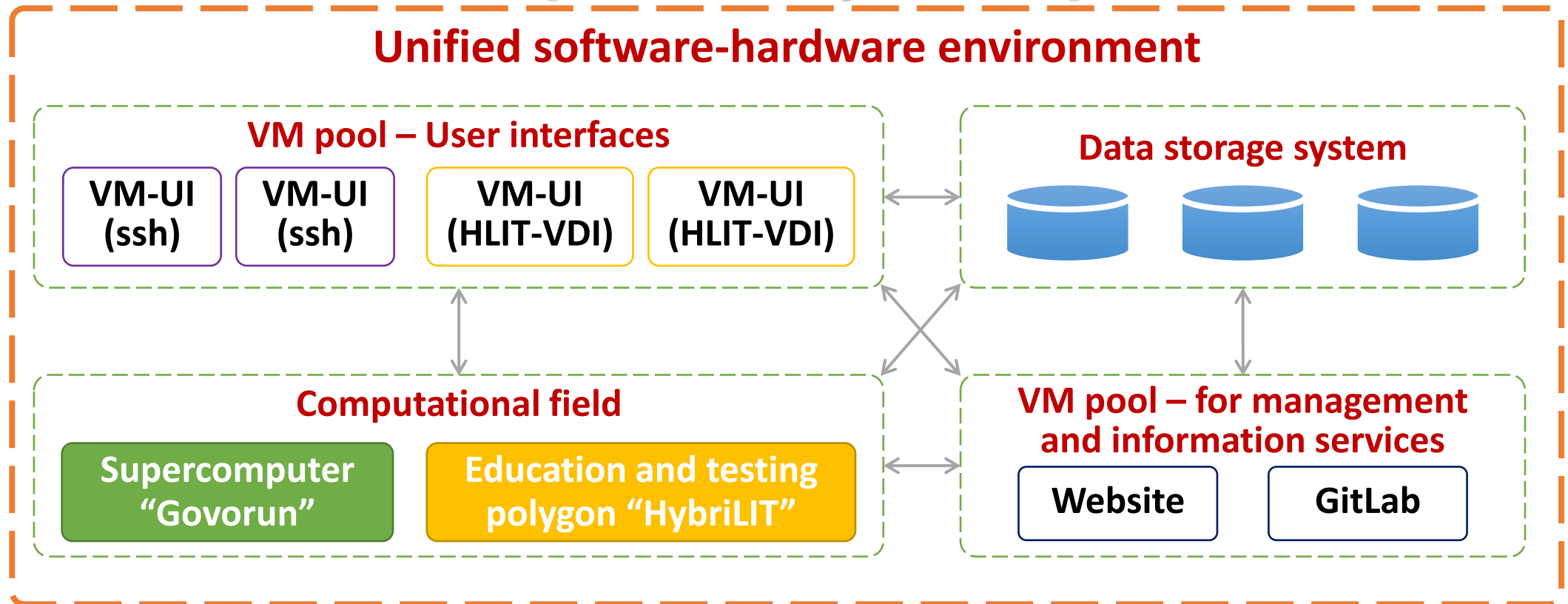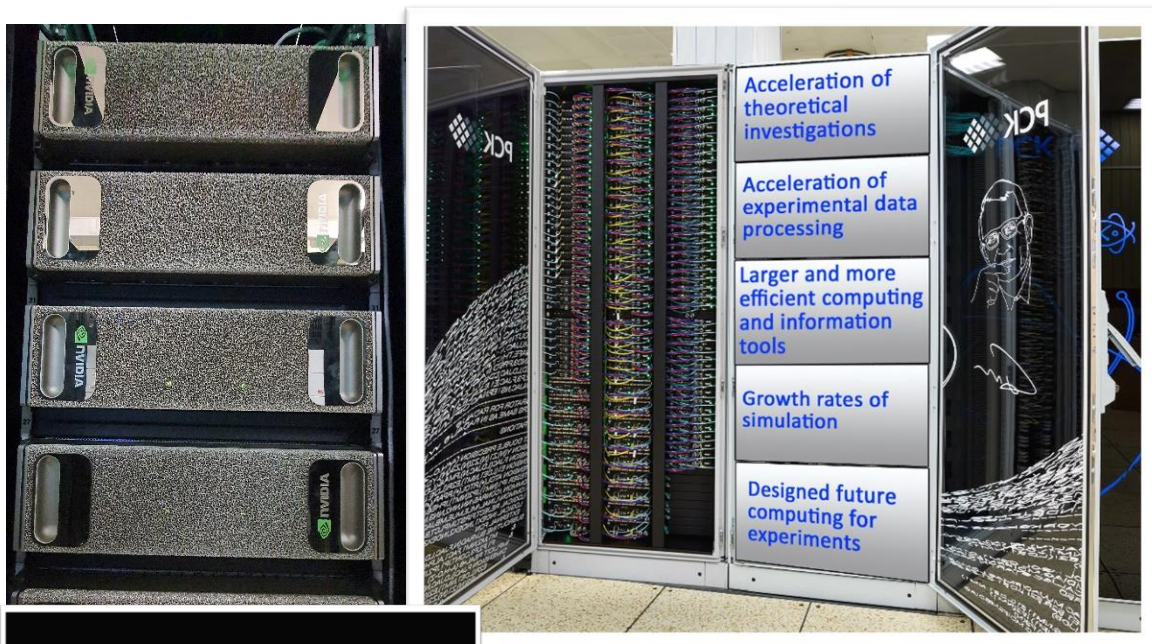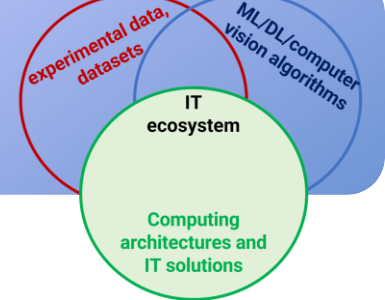The unified software and information environment of the HybriLIT platform allows users to use the education and testing polygon aimed at exploring the possibilities of novel computing architectures and IT solutions, to develop and debug their applications, furthermore, to carry out calculations on the supercomputer, which enables to effectively use the supercomputer resources.

IT ecosystem: computing architectures and IT solutions

"Govorun" Supercomputer

The "Govorun" supercomputer is a hyper-converged software-defined system, ranked **17th** in the current edition of the **IO500** list (July 2020). For the high-speed data storage system, RSC Group received the prestigious Russian DC Award 2020 in "The Best IT Solution for Data Centers" nomination at the awards ceremony held on 10 December 2020 in Moscow.

GPU component
based on NVIDIA DGX-1 Volta

CPU component
based on the latest Intel architectures:
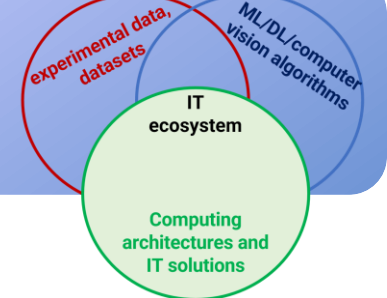Intel Xeon Phi gen.2 and Intel CascadeLake processors

Total peak performance:
1.7 PFLOPS SP
860 TFLOPS DP
300 Gb/s Data IO rate

IT ecosystem: computing architectures and IT solutions

GPU component of the "Govorun" Supercomputer

The GPU component consists of **5 NVIDIA DGX-1 servers**. Each server has **8 GPUs NVIDIA Tesla V100** based on the latest architecture NVIDIA Volta. Moreover, one server NVIDIA DGX-1 has **40,960** CUDA cores, which are equivalent to 800 high-performance central processors. A whole number of novel technologies are used in DGX-1, including the NVLink 2.0 wire with a bandwidth of up to 300 Gb/s.

The GPU component gives supercomputer users the ability to perform massively parallel computing for general-purpose tasks using such technologies as CUDA and OpenCL, as well as to use applications already adapted for this architecture. In addition, the GPU component allows using machine learning and deep learning algorithms to solve applied problems by a neural network approach.
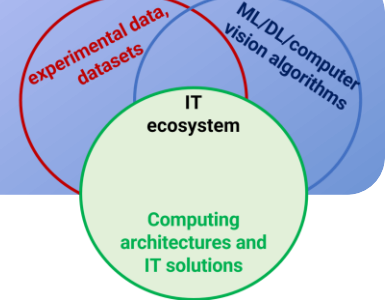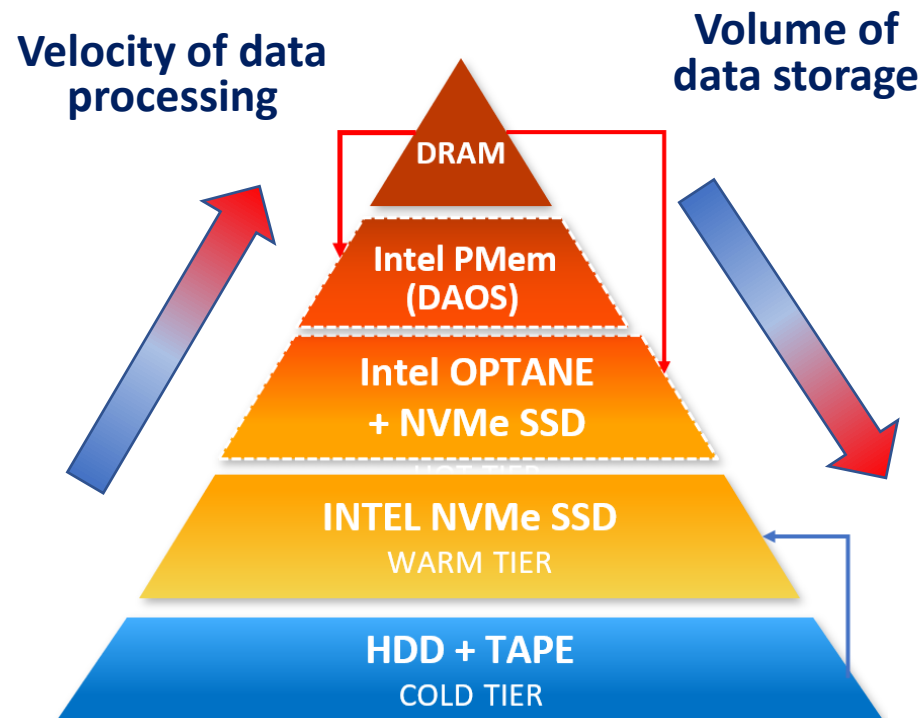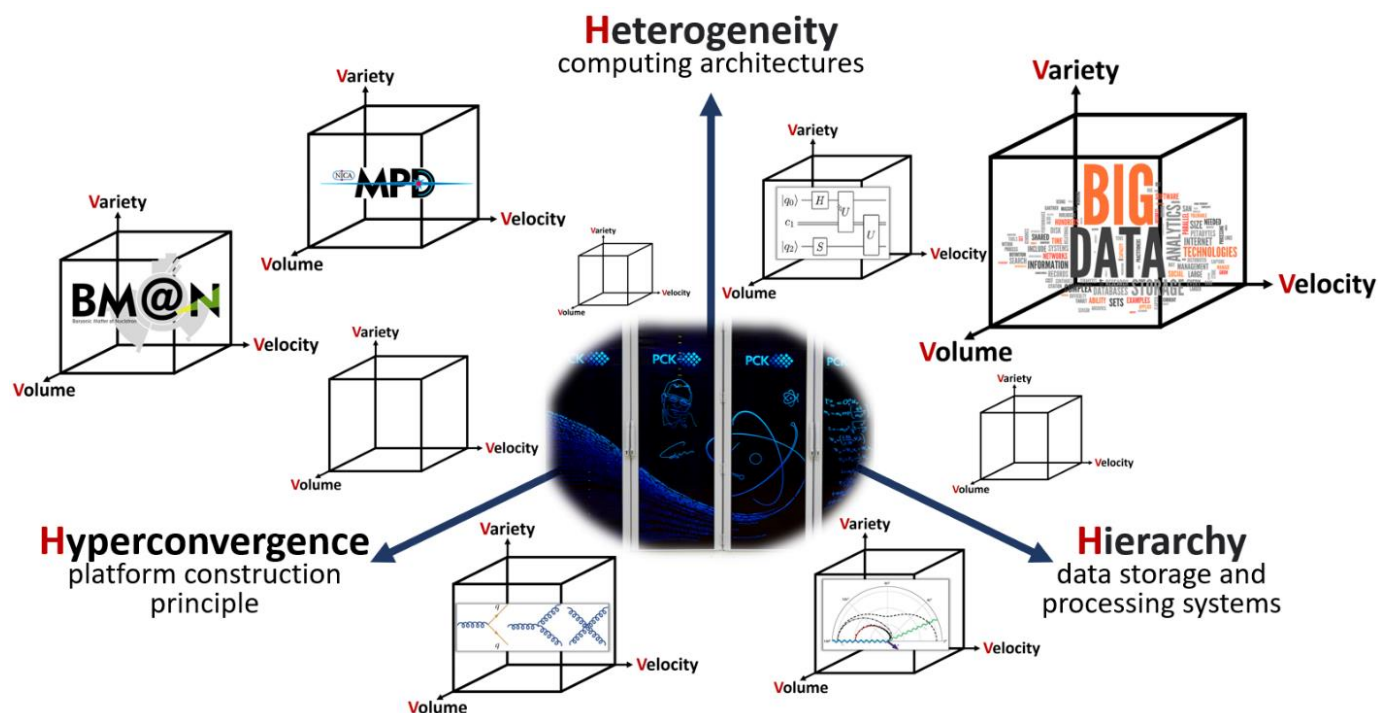
The DAOS polygon of the "Govorun" supercomputer takes the **1st** place among Russian supercomputers in terms of the data processing rate in the current **IO500 list**.
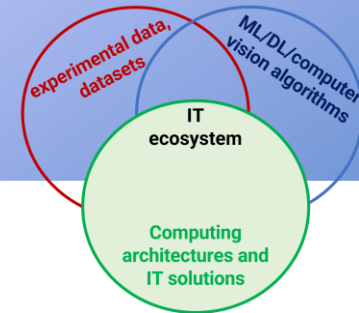
Heterogeneity
Hierarchy
Hyperconvergence
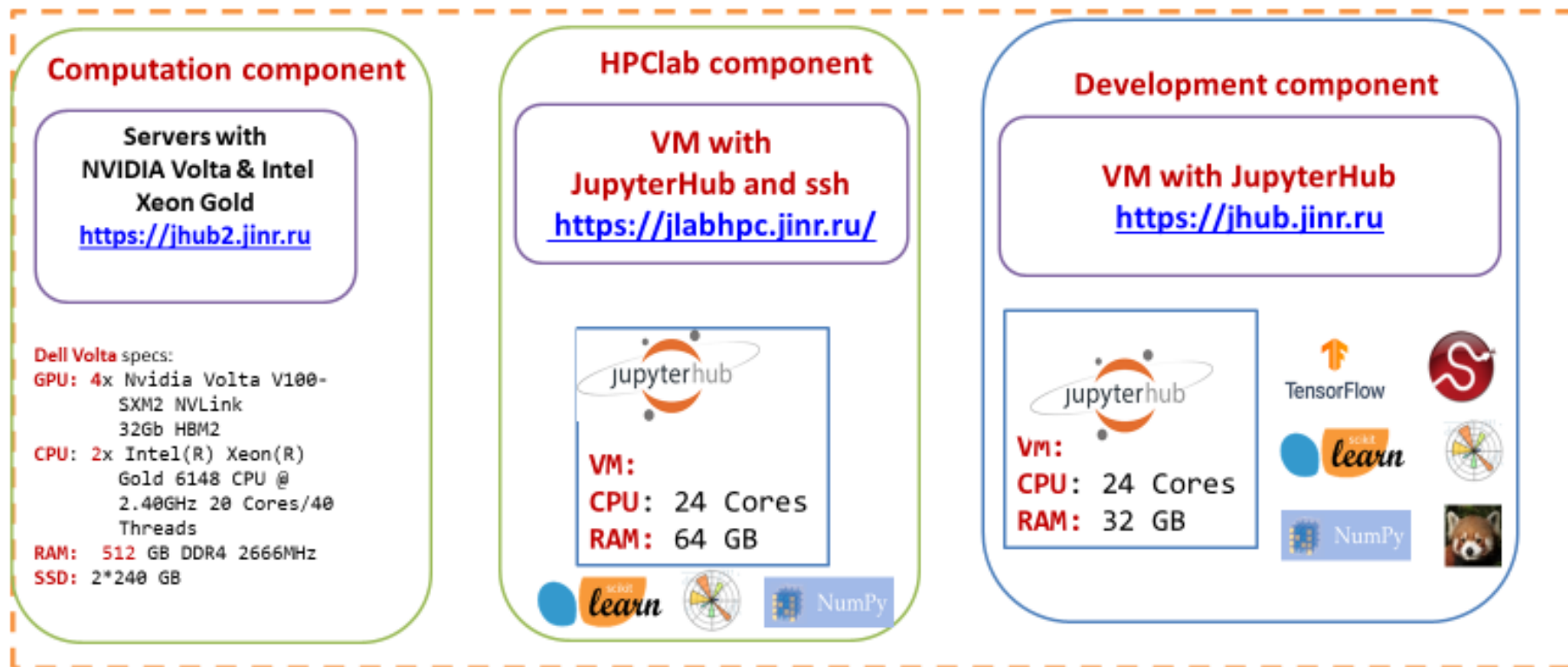
provide →

Variety
Velocity
Volume

Velocity of data processing

Volume of data storage

DRAM

Intel PMem (DAOS)

Intel OPTANE + NVMe SSD

INTEL NVMe SSD
WARM TIER

HDD + TAPE
COLD TIER

Environment for development

Ecosystem for HPC and ML/DL tasks

Computation component

Servers with NVIDIA Volta & Intel Xeon Gold
https://jhub2.jinr.ru

Dell Volta specs:
GPU: 4x Nvidia Volta V100-
       SXM2 NVLink
       32Gb HBM2
CPU: 2x Intel(R) Xeon(R)
       Gold 6148 CPU @
       2.40GHz 20 Cores/40
       Threads
RAM: 512 GB DDR4 2666MHz
SSD: 2*240 GB

HPClab component

VM with JupyterHub and ssh
https://jlabhpc.jinr.ru/

jupyterhub

VM:
CPU: 24 Cores
RAM: 64 GB

Development component

VM with JupyterHub
https://jhub.jinr.ru

jupyterhub

VM:
CPU: 24 Cores
RAM: 32 GB
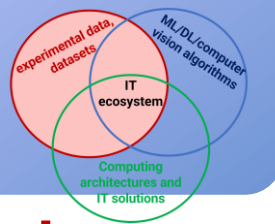
experimental data

labeled/ annotated data

Video files

data labeling/annotation tool
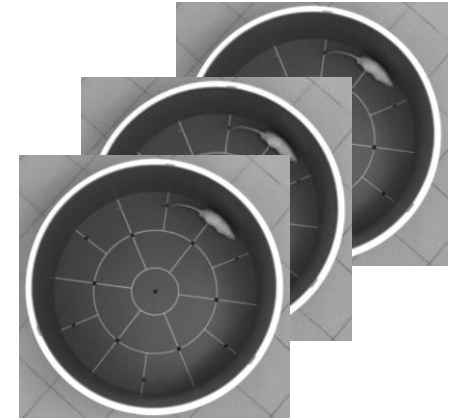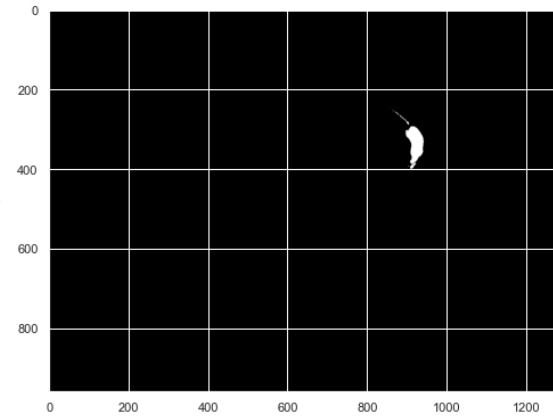
Action recognition

Image files
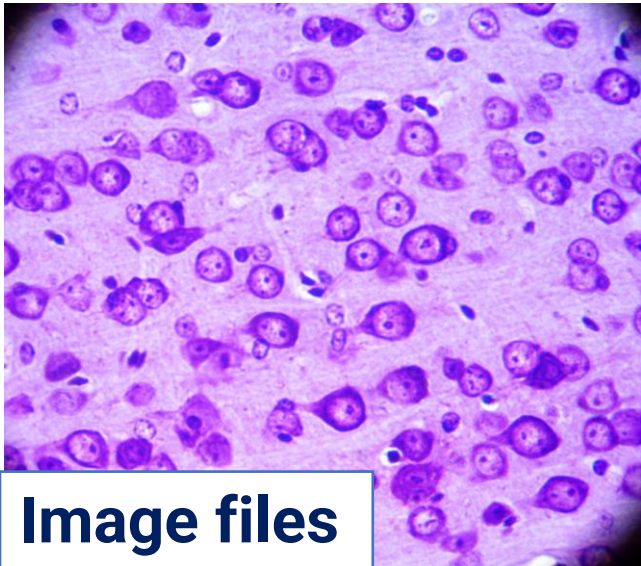
data labeling/annotation tool

Data labeling
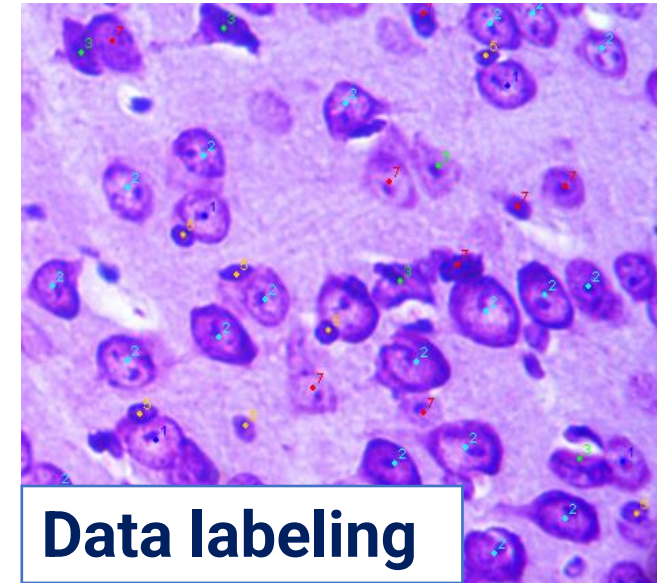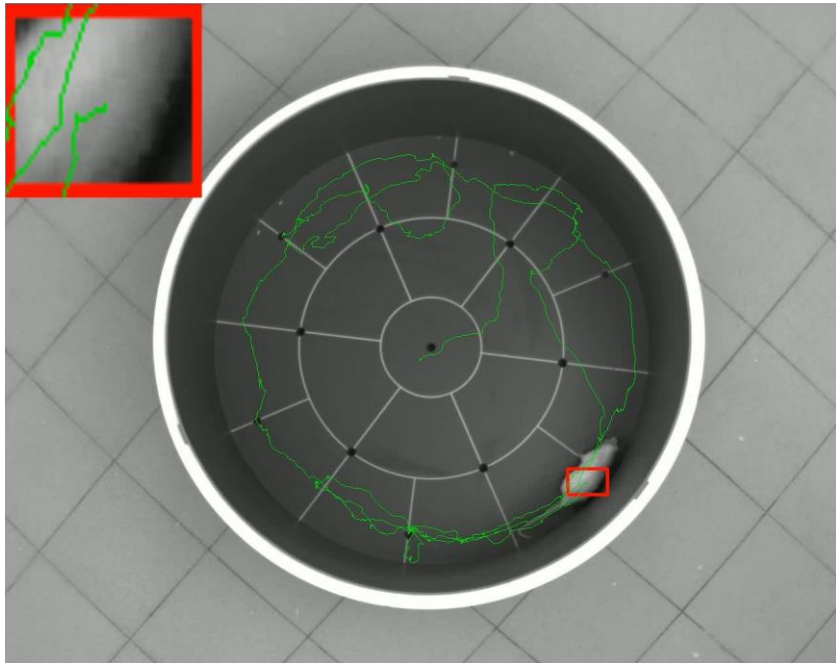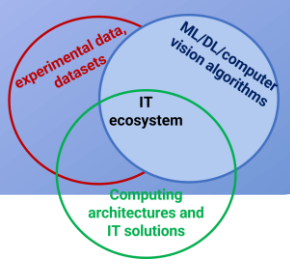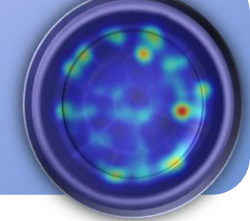
Tracking a laboratory animal:
Approaches:
- Tracking using a correlation filter
- Tracking with object-background segmentation assessment Best: CSRT tracker
- GOTURN
- MOSSE
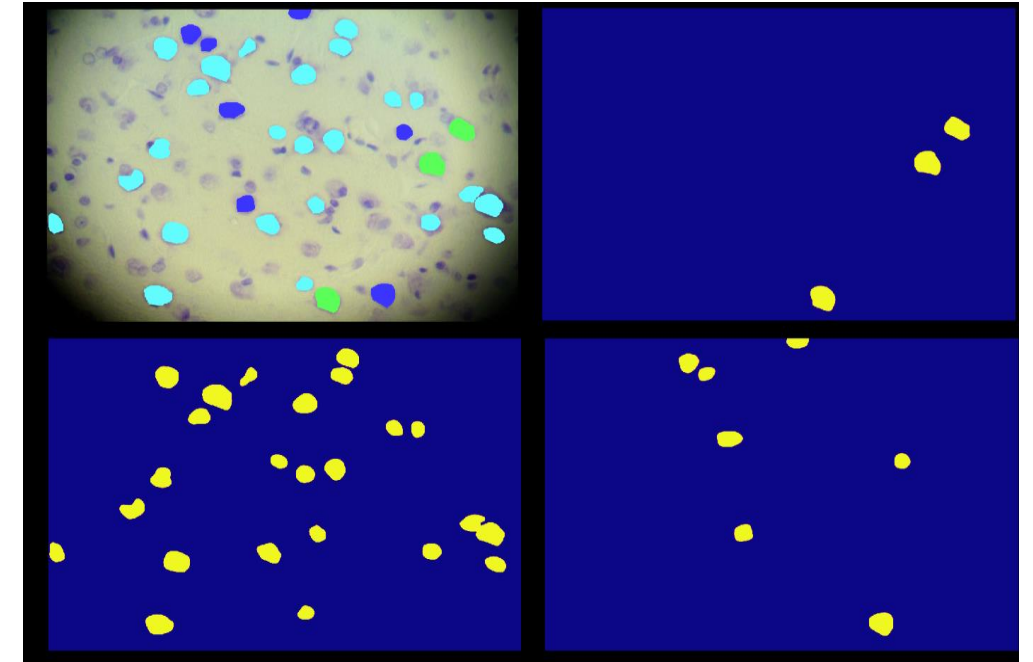
Neural networks for the task of neuron segmentation on brain slice images

IT ecosystem: web service

BIOHLIT

На главную   К экспериментам

Все эксперименты / вьетнам

Эксперимент:

🖉 Редактировать

вьетнам

Описание:

лечение

Дата забоя:	Тип животных:	Пол животных:
28-06-2022	mouse	

Дата завоза:	Дата рождения животных	Дата облучения:
28-06-2022	28-06-2022	28-06-2022

Облучение	Доза Gr	Другие воздействия
	5Гр	

BIOHLIT

На главную   К экспериментам   Пользователь: Oksana Streltsova   Выйти 🚪

Все эксперименты / Эксперимент / control

Поведение животных

| Открытое поле | Моррис | Т-лабиринт |

Загрузить

Всего видео:7

Выбраны:

Анализ группы

☐ №	Скачать	Имя файла	Статус	Запущен	Завершено			
☐ 1	Ссылка	Control_mouse1 10-22-30.avi	Running/Pending	2022-07-06 13:15:15		Анализ	Результаты	Удалить файл
☐ 2	Ссылка	Control_mouse2 10-31-24.avi	FAIL	2022-06-28 11:25:36	2022-06-28 11:25:52	Анализ	Результаты	Удалить файл
☐ 3	Ссылка	Control_mouse3 10-38-39.avi	FAIL	2022-06-28 11:27:16	2022-06-28 11:27:33	Анализ	Результаты	Удалить файл

IT ecosystem: web service
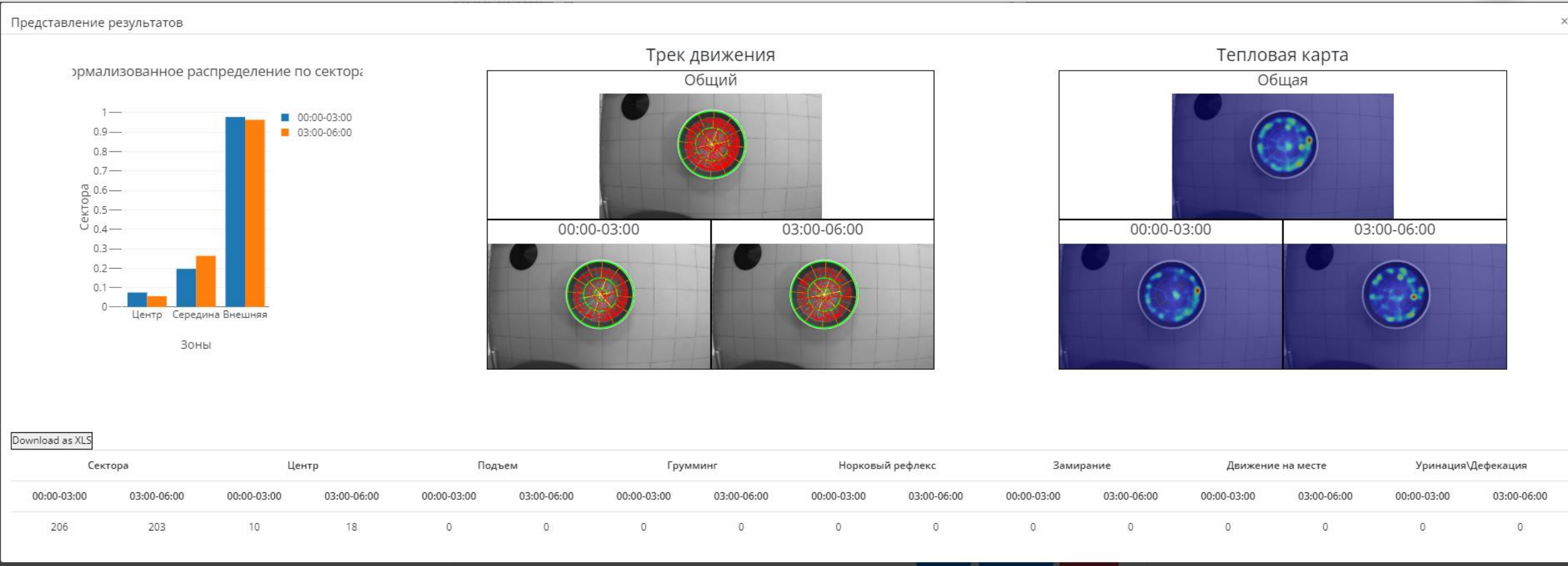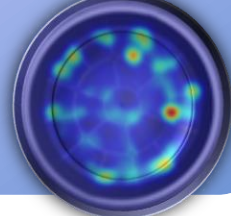
IT ecosystem
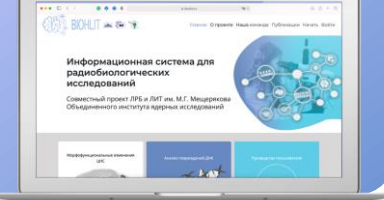
Computational field
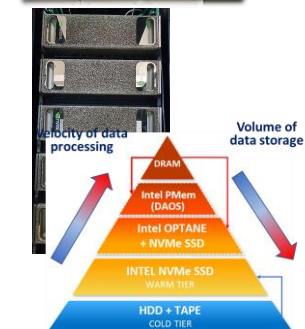
**Tiered storage system**

Databases
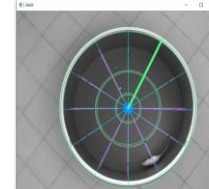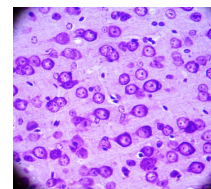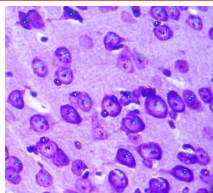
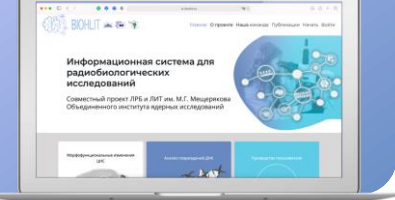**Data markup and annotation services**

Algorithmic block for behavioral analytics

Algorithmic block for histological studies

Algorithmic block for a comprehensive study of morphofunctional changes

Web services for data loading, data analysis

The studies are carried out using all the capabilities of the HybriLIT platform
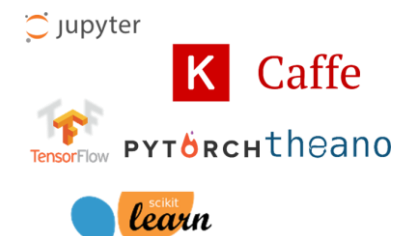
The **joint project of MLIT and LRB** is focused on creating an Information System (**IS**) for analyzing behavioral and pathomorphological changes in the central nervous system when studying the effects of ionizing radiation and other factors on biological objects.
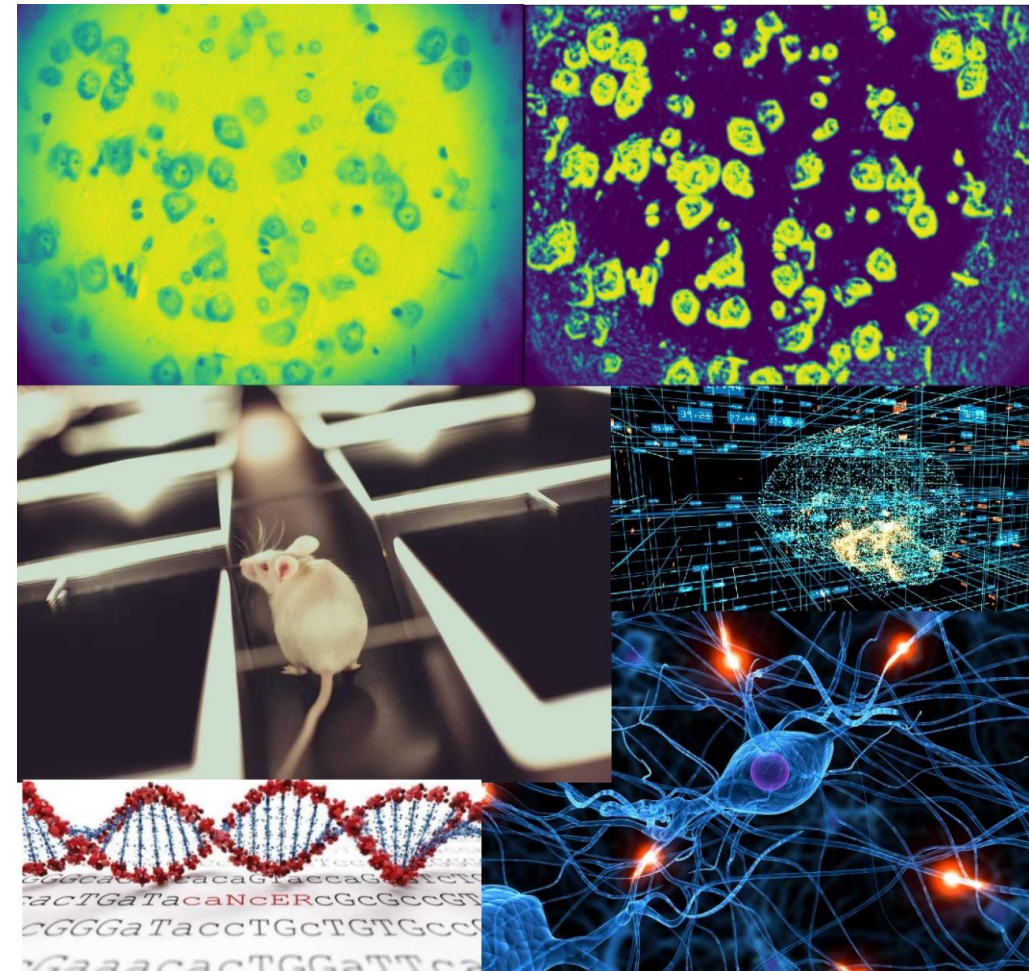
The **IS** is based on:

- computer vision algorithms on top of machine and deep learning technologies (ML/DL);
- modern IT solutions for data storage, processing and visualization.

The **IS** will allow one to simplify and accelerate:

- experimental data processing by automating the morphological classification of neural cells;

- data analysis techniques using the latest neural network algorithms based on ML/DL;

- work with experimental data for different research groups;

- experimental data systematization and development of effective methods for preventing and countering the negative effect of ionizing radiation.

Thank you for your attention!