

Отчет за 2024 год

о выполнении работ,
финансируемых грантом молодых
ученых и специалистов ОИЯИ
по конкурсу:
«МОЛОДОЙ НАУЧНЫЙ СОТРУДНИК»

ГРАНТОПОЛУЧАТЕЛЬ:

ПРЯХИНА ДАРЬЯ ИГОРЕВНА

НТО ВКиРИС
НАУЧНЫЙ СОТРУДНИК



Объявленные задачи проекта



I. Провести тестирование и доработать программный комплекс для создания цифровых двойников (ЦД) распределенных центров сбора, хранения и обработки данных (РЦОД).

II. Создать ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

III. Подготовить кандидатскую диссертацию.

IV. Подключить систему авторизации SSO и разместить в общий доступ.



Объявленные задачи проекта

I. Провести тестирование и доработать программный комплекс для создания цифровых двойников (ЦД) распределенных центров сбора, хранения и обработки данных (РЦОД).

ВЫПОЛНЕНО

II. Создать ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

ВЫПОЛНЕНО

III. Подготовить кандидатскую диссертацию.

ВЫПОЛНЕНО

IV. Подключить систему авторизации SSO и разместить в общий доступ.

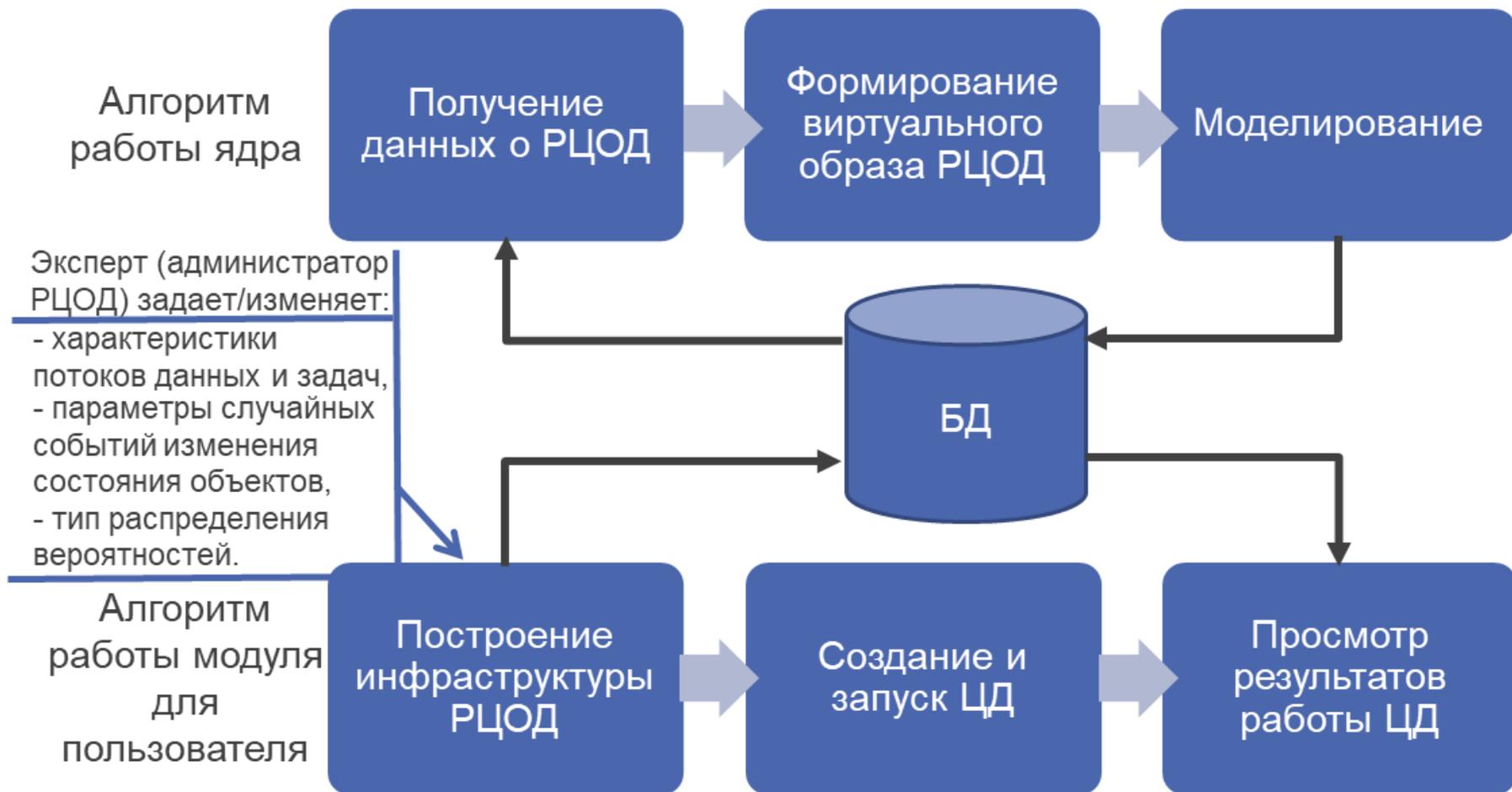
НЕ ВЫПОЛНЕНО

Выполненные работы

I. Проведено тестирование и доработан программный комплекс для создания цифровых двойников РЦОД.

Состав программного комплекса:

- база данных (БД);
- ядро ЦД (программа моделирования);
- веб-сервис (модуль для пользователя).



Выполненные работы

I. Проведено тестирование и доработан программный комплекс для создания цифровых двойников РЦОД.

Благодарность за помощь в выполнении работы Владимир Валентинович Трофимов (ЛИТ)

- тестирование программы;
- консультации и полезные советы;
- новые идеи и интересные предложения;
- участие в подготовке публикаций.



Публикация результатов

➤ Priakhina D., Korenkov V., Trofimov V.
Prototype of a software complex for creating digital twins of distributed data acquisition, storage and processing centers



ЭЧАЯ

Физика
элементарных частиц и
атомного ядра

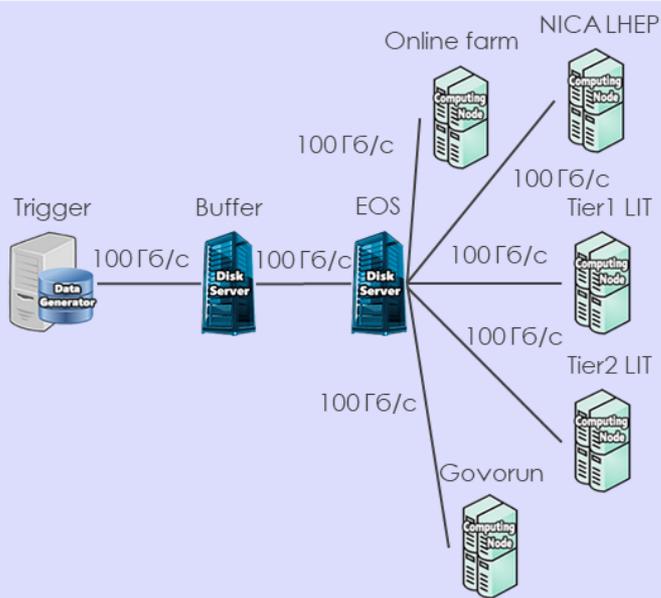
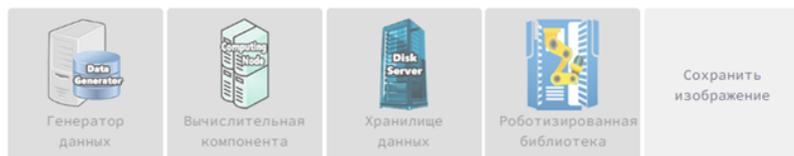
Том 55, №3, стр. 444-446 (2024)

Выполненные работы

II. Создание ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

Применение программного комплекса для построения ЦД вычислительной инфраструктуры эксперимента VM@N

Построение инфраструктуры центра сбора, хранения и обработки данных



Цели:

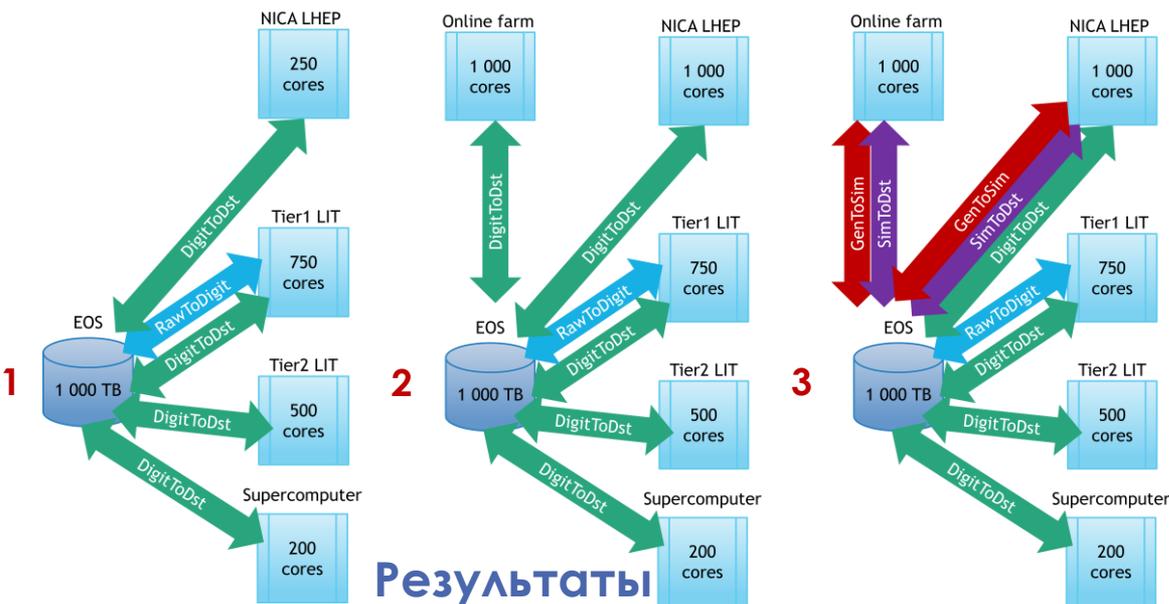
- оценить потребности в ресурсах для хранения и обработки данных с учетом планируемых параметров потоков данных будущих сеансов эксперимента;
- получить примерное время, которое потребуется на обработку данных по окончании сеанса эксперимента.

Выполненные работы

II. Создание ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

Применение программного комплекса для построения ЦД вычислительной инфраструктуры эксперимента VM@N

Сравнение конфигураций инфраструктуры для обработки данных



1 → 2:

– Увеличено количество ресурсов **NICA LHEP** с 250 до 1 000 ядер.

– Добавлены ресурсы **Online farm (1 000 ядер)**.

3: Добавлены потоки задач по обработке модельных данных.

100% использование ресурсов

Процесс преобразования raw-данных **ускорен в 2 раза**

	1	2	3
Преобразование всех raw-данных в формат digit (RawToDigit задачи)	60 часов (2,5 дня)	28 часов (1 день)	28 часов (1 день)
Время полной обработки всех экспериментальных данных	432 часов (18 дней)	240 часов (10 дней)	367 часов (15 дней)
Время полной обработки всех модельных данных	---	---	130 часов (6 дней)

Выполненные работы

II. Создание ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

Применение программного комплекса для построения ЦД вычислительной инфраструктуры эксперимента VM@N

Благодарность за помощь в выполнении работы

Константин Герценбергер (ЛФВЭ)

- предоставление необходимой информации об эксперименте;
- предоставление исходных данных для моделирования;
- консультации, советы и интересные вопросы;
- участие в подготовке публикаций;
- проявленный интерес к проекту.



Публикация результатов

➤ **Пряхина Д.И., Кореньков В.В., Трофимов В.В., Герценбергер К.В.**
Верификация программы моделирования для создания цифровых двойников распределенных центров сбора, хранения и обработки данных

International Journal of Open Information Technologies Том 12, №1, стр. 118-128 (2024)

Выполненные работы

II. Создание ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

Применение программного комплекса для построения ЦД вычислительной системы онлайн-фильтра данных эксперимента SPD

Построение инфраструктуры центра сбора, хранения и обработки данных



Генератор данных



Вычислительная компонента



Хранилище данных



Роботизированная библиотека

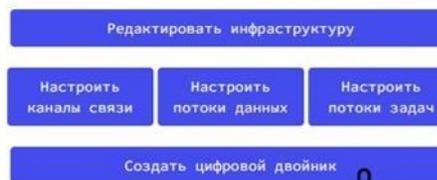
Сохранить изображение

Trigger

Buffer

Computing

Intermediate



Цели:

- спроектировать эффективную вычислительную систему для обеспечения быстрой реконструкции событий и фильтрации данных в режиме реального времени;
- предусмотреть возможности для дальнейшего развития и оптимизации проектируемой системы.

Выполненные работы

II. Создание ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

Применение программного комплекса для построения ЦД вычислительной системы онлайн-фильтра данных эксперимента SPD

Исходные данные		Результаты	
Время работы установки	24 ч	Буфер для хранения raw-данных	1 100 ТБ
Скорость генерации данных	20 ГБ/с	Промежуточное хранилище	3 100 ТБ
Эффективность генерации данных	60%	Вычислительные ресурсы	7 000 ядер
Скорость обработки событий	1 000 соб/с	Канал связи Триггер-Буфер	150 Гб/с
Количество задач обработки данных	100 000	Канал связи Буфер-Выч.рес.	80 Гб/с
Объем raw-файла	4 ГБ	Канал связи Выч.рес.-Хранилище	500 Гб/с
Объем dec-файла	4 ГБ		
Объем pres-файла	8 ГБ		
Объем filtred-файла	450 МБ		

1. Получены оценки по требуемому оборудованию для решения функциональных задач онлайн-фильтра в режиме реального времени с учетом планируемых характеристик потоков данных эксперимента.

2. Результаты доказали возможность использования СПО для решения задачи проектирования РЦОД.

Выполненные работы

II. Создание ЦД вычислительных инфраструктур для экспериментов комплекса NICA.

Применение программного комплекса для построения ЦД вычислительной системы онлайн-фильтра данных эксперимента SPD

Благодарность за помощь в выполнении работы

Данила Олейник (ЛИТ)

- предоставление необходимой информации об эксперименте;
- предоставление исходных данных для моделирования;
- консультации при подготовке докладов;
- участие в подготовке публикации;
- проявленный интерес к проекту.



Публикация результатов

➤ **Пряхина Д.И., Кореньков В.В., Олейник Д.А., Жемчугов А.С. Цифровой двойник системы первичной обработки данных эксперимента SPD (отправлено в журнал)**



Выполненные работы

II. Создание ЦД
вычислительных инфраструктур
для экспериментов комплекса
NICA.

Представление
результатов

59th meeting of the PAC for Particle Physics
22 January, 2024, JINR, Dubna, Russia

AYSS 28.10-01.11 2024
28th International Scientific Conference of Young Scientists and Specialists
Digital twins for solving management and development tasks of distributed computer systems for megascience projects
DARIA PRIAKHINA
Meshcheryakov Laboratory of Information Technologies
Joint Institute for Nuclear Research
28.10.2024

28th International Scientific Conference
of Young Scientists and Specialists
28.10-01.11, 2024, JINR, Dubna, Russia

Приглашенный секционный доклад

Software complex for creating digital twins of large-scale distributed computer systems for megascience projects
D. Priakhina, V. Korenkov, V. Trofimov
priahinad@jinr.ru

Digital Twin (DT) – a computer model that can be used in real time throughout the entire life cycle of distributed data acquisition, storage and processing center (DDC) [1, 2].

Software complex architecture

- Web-service
- equipment configuration
- data flow parameters
- simulation results
- base core for transferring and processing data modeling

Functional purposes of digital twins

- Designing a DDC.
- Analyzing the efficiency and reliability of DDC functioning.
- Testing scaling scenarios taking into account the requirements for data flows and job flows.
- Assessment of the required amount of resources for specific tasks.
- Checking strategies for managing job flows.

Distinctive features of modeling core

- Universal – applicable for modeling any data center without changing the program code.
- The initial data: the architecture and hardware parameters of the DDC-equipment; the characteristics of data flows and job flows.
- Probabilistic distributions are taken into account when forming data flows, job flows, and criteria for the functioning of equipment.
- Used for design tasks, data center scaling during operation, searching for problem areas when data flows and jobs flows change.

Stages of using software complex:

1. Building an infrastructure for data acquisition, storage and processing centers.
2. Configuring the simulation parameters: setting the duration of the DDC operation; adding probabilistic events that may occur in the system (equipment failure, changes in the amount of computing resources, etc.); adding specific objects and events for logging; setting additional possible modifications besides the basic hardware configuration that was set when creating the DDC infrastructure.
3. DT launch.
4. Analysis of the DT functioning results: data storages load; using computing resources; load rate; job queues; completed jobs; distribution of files in storages.

Implementation tools: CORE, WEB, JABBY, GEMMA, etc.

Software complex application: Digital Twin of BMIN computing infrastructure for data production; Comparison of different computing infrastructure configurations for data processing; More detailed information is available in [3, 4].

References

1. Priakhina D., Korenkov V. The relevance of creating a digital twin for managing distributed data acquisition, storage and processing centers // Modern Information Technologies and IT Education. 2023. V. 19, no. 3 (in Russ.).
2. Priakhina D., Korenkov V., Trofimov V. A method of constructing digital twins for solving problems of effective management and development of distributed data acquisition, storage and processing centers // Modern Information Technologies and IT Education. 2023. V. 19, no. 3 (in Russ.).
3. Priakhina D., Korenkov V., Trofimov V., Gertenberger E. Verification of the simulation program for creating digital twins of distributed data acquisition, storage and processing centers // International Journal of Open Information Technologies. 2024. V. 12, no. 1. P. 118-128 (in Russ.).
4. Priakhina D., Korenkov V., Trofimov V., Gertenberger E. Simulation Results of BMIN Computing Infrastructure // Physics of Particles and Related Letters. 2023. V. 20, no. 6. P. 1272-1275.

Выполненные работы

III. Подготовка кандидатской диссертации.

Защита состоялась **5 декабря 2024 г.**
в Санкт-Петербургском
государственном университете

Ученая степень:

кандидат технических наук

Научная специальность:

2.3.1. Системный анализ,
управление и обработка
информации, статистика

Тема:

«Цифровые двойники для
решения задач управления и развития
распределенных центров сбора,
хранения и обработки данных»

Научный руководитель:

д.т.н. Кореньков Владимир Васильевич



Дополнительная деятельность



Выполненные работы (дополнительно)

Образовательная активность

«Подготовка специалистов в области вычислительной физики и информационных технологий»





ЛАБОРАТОРИЯ
ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ
имени М.Г. Мещерякова

Help desk Контакты Сборщик RSS-лент English Search

О ЛИТ СТРУКТУРА НАУЧНАЯ ЖИЗНЬ ИТ - ИНФРАСТРУКТУРА ОБРАЗОВАНИЕ СОБЫТИЯ И ДАТЫ INFOBOARD

Об образовательной активности



Поддержка раздела «Образование» на сайте ЛИТ при участии Лидии Анатольевны Калмыковой

Участники активности

Подготовка специалистов в области вычислительной физики и информационных технологий

— Активность в рамках темы 1119

Цель

Подготовка и переподготовка специалистов в области вычислительной физики и информационных технологий на базе Многофункционального информационно-вычислительного комплекса (МИВК) Объединенного института ядерных исследований (ОИЯИ) и его учебно-образовательных компонент осуществляются в целях:

- повышения квалификации сотрудников ОИЯИ для развития научных проектов, в том числе класса мегасайнс, реализуемых в ОИЯИ или с его участием, а также для создания и поддержки цифровой экосистемы (ЦЭС) ОИЯИ;
- распространения компетенций в области вычислительной физики и информационных технологий в регионы России и страны-участницы ОИЯИ для увеличения кадрового потенциала ОИЯИ и сотрудничающих с Институтом организаций.

Обоснование

Основной предпосылкой к созданию активности является необходимость формирования научно-исследовательской среды для обеспечения профессионального роста ИТ-специалистов, создание и развитие научных групп, привлечение новых сотрудников в проекты ОИЯИ. Дополнительная подготовка кадров преимущественно по заказу лабораторий ОИЯИ должна быть направлена на развитие специальных компетенций, углубленных знаний и навыков практического характера в области вычислительной физики и информационных технологий.

Актуальность привлечения студентов связана с возможностью их участия на протяжении длительного отрезка времени в реальных проектах. Начиная с учебных практик студенты могут решать небольшие задачи, приобретая опыт работы в команде и изучая детали проекта. Далее наиболее перспективные студенты получат возможность продолжить работу в научных группах, подготовить бакалаврские работы, магистерские диссертации. В дальнейшем при поступлении в аспирантуру и смогут стать стажерами-исследователями ОИЯИ.

Для решения проблемы привлечения студентов предлагается разработать учебные курсы, организовать мероприятия (школы, конференции, хакатоны и др.) на базе Информационных центров ОИЯИ, а также университетов, сотрудничающих с ОИЯИ.

Задачи



Владимир Васильевич Корнев
научный руководитель



Дарья Пряхина
руководитель



Оксана Стрельцова
руководитель

Выполненные работы (дополнительно)

Образовательная активность

«Подготовка специалистов в области вычислительной физики и информационных технологий»



Образовательная активность в ЛИТ ОИЯИ 2019 – 2023



Сбор информации от сотрудников ЛИТ по работе со студентами



За 2024 г. собиралась дополнительная информация о совместных публикациях и выступлениях на конференциях (студент + науч. рук.)



Наиболее активные сотрудники ЛИТ, вовлеченные в руководство квалификационными работами студентов

Выполненные работы (дополнительно)



Образовательная активность

«Подготовка специалистов в области вычислительной физики и информационных технологий»

 **Весенняя Школа**  **2024** 
по информационным технологиям ОИЯИ

15 - 16 Апреля



Конкурс на лучший доклад
10 победителей

Студенческие отчеты в виде презентаций с результатами проделанной работы и планами на дальнейшее развитие проекта.

30 студентов из 7 университетов РФ



Выполненные работы (дополнительно)

Образовательная активность

«Подготовка специалистов в области вычислительной физики
и информационных технологий»



**Результаты проведения ИТ-Школ за 2 учебных года
(2022/23 и 2023/24)
по состоянию на 1.12.2024**

117

студентов
участвовали в
ИТ-школах ОИЯИ

30

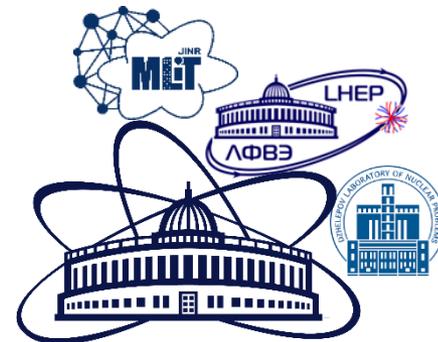
дипломных работ
бакалавров и
магистров

20

студентов не выпускных курсов
ведут работы над проектами ОИЯИ

18

стали
сотрудниками ОИЯИ



Выполненные работы (дополнительно)



Образовательная активность

«Подготовка специалистов в области вычислительной физики и информационных технологий»

IT SCHOOL
JINR



Осенняя Школа

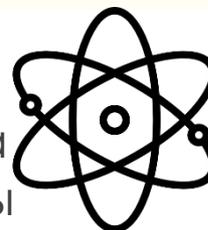
по информационным технологиям ОИЯИ

7 - 11 Октября 2024



58 студентов

от 2 курса бакалавриата
до 2 курса магистратуры



6 научных направлений

более 20 проектов для
выбора темы ВКР



11 вузов РФ

от Камчатки
до Санкт-Петербурга



более 30 лекторов

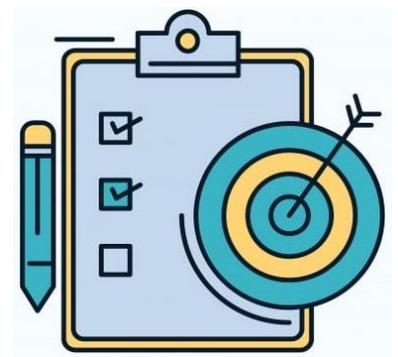
5 организаций
6 подразделений ОИЯИ

Результат на 1.12.24

38 студентов

- определились с направлением;
- начали общение с руководителями.

ИТОГИ 2024 года



1. Количество публикаций: 4

3 опубликовано + 1 принято в печать

2. Количество научных мероприятий: 3 (лично)

3. Участие в орг. комитетах научных мероприятий: 2 (ученый секретарь)

4. Помощь в подготовке к лицензированию образовательной программы для Филиала МГУ в г. Дубне,

в том числе учебного плана и рабочих программ дисциплин по направлению 01.04.02 Прикладная математика и информатика (профиль «Методы и технологии обработки данных в гетерогенных вычислительных средах»)



Планы на 2025 год

Научная деятельность

1. Доработка программного комплекса для создания цифровых двойников РЦОД:
 - по результатам тестирования;
 - по требованиям всех заинтересованных.
2. Подключение системы авторизации SSO.
3. Размещение программного комплекса в общем доступе для сотрудников ОИЯИ, предоставление возможности его самостоятельного использования для создания цифровых двойников.
4. Применение программного комплекса для создания цифровых двойников вычислительных инфраструктур комплекса NICA и других.

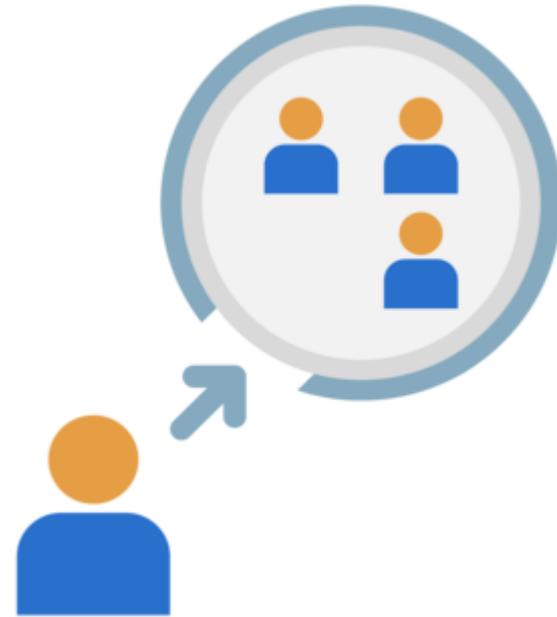


Планы на 2025 год

Образовательная активность

1. Разработка информационной системы для образовательной активности.
2. Разработка сайта для Школы по информационным технологиям ОИЯИ.
3. Разработка модуля для сбора, хранения и анализа данных о результатах образовательной активности.
4. Проведение Школ, семинаров и пр.

Готова приступить
К НОВЫМ ЗАДАЧАМ
и присоединиться
к проектам



**СПАСИБО
ЗА ВНИМАНИЕ !**

