



**Федеральное государственное автономное  
образовательное учреждение высшего образования  
«Национальный исследовательский университет «Высшая школа экономики»  
(НИУ ВШЭ)**

**ПРОТОКОЛ № 8.3.1.8-07/15/22  
электронного голосования ученого совета НИУ ВШЭ – Санкт-Петербург  
от 28.12.2022**

**Председатель** – Тыщецкая А.Ю.

**Ученый секретарь** – Гигаури Д.И.

В электронном голосовании приняли участие: Абрамов В.В., Гордин В.Э., Декальчук А.А., Жуков А.Е., Заривной Н.А., Зеленев Е.И., Кузнецов А.М., Кузьмин П.В., Лимонов Л.Э., Мифтахутдинов Р.И., Муравьев А.А., Мухин М.С., Немировская А.В., Омельченко Е.Л., Плахотник М.С., Почекаев Р.Ю., Сивицкий В.А., Стародубцев А.В., Токарев Д.В., Ходачек А.М., Цветкова Л.А., Яровенко К.В.

**Кворум имеется, голосование правомочно.**

**1. О создании в структуре факультета Санкт-Петербургская школа гуманитарных наук и искусств Лаборатории языковой конвергенции**

**ПОСТАНОВИЛИ:**

1.1. Одобрить концепцию лаборатории языковой конвергенции факультета Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург (приложение 1);

1.2. Одобрить создание в структуре факультета Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург лаборатории языковой конвергенции.

В процедуре электронного голосования приняли участие 22 членов ученого совета: «за» - 21; «воздержались» - 1; «против» - 0.

**2. О создании в структуре факультета Санкт-Петербургская школа физико-математических и компьютерных наук научной лаборатории Интернета вещей и киберфизических систем**

2.1. Одобрить концепцию научной лаборатории Интернета вещей и киберфизических систем факультета Санкт-Петербургская школа физико-математических и компьютерных наук (приложение 2);

2.2. Одобрить создание в структуре факультета Санкт-Петербургская школа физико-математических и компьютерных наук научной лаборатории Интернета вещей и киберфизических систем.

В процедуре электронного голосования приняли участие 22 членов ученого совета:  
«за» - 21; «воздержались» - 1; «против» - 0.

**Все вопросы повестки дня электронного голосования рассмотрены.**

Подсчет голосов по вопросам, включенным в повестку электронного голосования проводил Гигаури Д.И.

Председатель



А.Ю. Тыщецкая

Ученый секретарь



Д.И. Гигаури



**Концепция  
лаборатории языковой конвергенции факультета Санкт-Петербургская школа  
гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург**

**1. Предпосылки создания.**

Факультет Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ – Санкт-Петербург (далее – СПбШГНиИ) выполняет научные исследования в области истории, филологии и лингвистики. В рамках последних двух направлений ведется активная работа по исследованию текстовых коллекций и массивов цифровыми инструментами: корпусными методами, технологиями автоматической обработки текстов, при помощи моделей NLP. За последние несколько лет под руководством Т. Ю. Шерстиновой на базе СПбШГНиИ было выполнено несколько проектов в данной сфере: создан корпус русского рассказа первой трети XX в., на его основе проведены статистические исследования языка и стиля малой русской прозы, тематики и художественных особенностей русской литературы, проводилось описание прагматического аспекта повседневной речевой коммуникации, изучение социолингвистической вариативности и коммуникативного разнообразия русского повседневного общения. С 2021 г. под руководством Т. Ю. Шерстиновой успешно работает научно-учебная группа (НУГ) по теме «Методы искусственного интеллекта в филологических исследованиях», целью которой является изучение применимости методов машинного обучения к текстам художественной прозы ([https://spb.hse.ru/humart/complit/ai\\_philology/](https://spb.hse.ru/humart/complit/ai_philology/)). В работе НУГ принимают активное участие как студенты, так и магистранты департамента филологии СПбШГНиИ, а также приглашенные ведущие специалисты из других университетских центров в области искусственного интеллекта и ДН. В связи с открытием в 2021 году в СПбШГНиИ магистерской программы в области фундаментальной и прикладной лингвистики «Языковые технологии в бизнесе и образовании» существенно увеличился спектр разрабатываемых проблем, расширилась палитра используемых методов искусственного интеллекта. На сегодняшний момент под руководством преподавателей департамента, а также приглашенных руководителей из IT-компаний и исследовательских лабораторий выполняются филологические и лингвистические проекты фундаментальной и прикладной направленности: стилиметрический анализ русской прозы, изучение устной речевой коммуникации и языка повседневного общения, изучение русской литературы с помощью цифровых методов, разработка веб-приложений и моделей для контролируемой генерации текстов разных жанров и для различных таргет-групп читателей, распознавания речи в ситуации специфического профессионального общения; оптимизация маркетинговых стратегий на основе данных сентимент-анализа отзывов на продукт и т.д. В перспективе развития СПбШГНиИ — открытие аспирантуры по специальности 5.9.8. «Теоретическая, сопоставительная и прикладная лингвистика» и аспирантуры по специальности 5.9.1 «Русская литература». Создание собственного научно-исследовательского подразделения, способного объединить существующие многочисленные исследования, существенно расширит возможности научных



коллективов НИУ ВШЭ – Санкт-Петербург участвовать в грантовых конкурсах на проведение полномасштабных научных исследований.

## **2. Состояние предметной области в России и за рубежом.**

Исследование текстовой среды как источника Больших Данных предполагает получение сведений об обществе на различных временных этапах его существования (доминирующих эмоциях, смене фокуса мнений, изменений в ценностной картине мира, точек “географической” притягательности для миграции и т.д.) на основе количественного анализа текстов, написанных на естественном языке. Данная проблематика отчасти пересекается с точками интереса так называемой Цифровой гуманитаристики. В России ведущие ДН-центры локализованы в ВШЭ (московский и пермский кампусы), за рубежом - в Городском университете Нью-Йорка (США), в университете Бергена (Норвегия), в университете Умеа (Швеция) и др.

Другая точка пересечения - использование методов машинного обучения для извлечения различной информации из текстовых данных: распознавание сущностей, географических названий с последующим их наложением на карты; тематическое моделирование и т.д. Стейкхолдерами в данной области в России являются научно-исследовательские лаборатории в МФТИ, СПбГУ, ТГУ и ПГНИУ, за рубежом - MIT, Berkley University, ряд крупных университетов Гонконга и Китая.

## **3. Обоснование создания.**

Факультету Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ – Санкт-Петербург необходима собственная структурная единица для организации и проведения исследования больших массивов текстовых данных методами искусственного интеллекта по нескольким причинам.

1. Для формирования единого научного коллектива, способного решать амбициозные и актуальные задачи в области NLP (обработка естественного языка), а также тесно связанной с ней области — Natural Language Understanding (понимание естественного языка), проецируя уже существующие в них модели и технологии в сферу таких массивов текстовых данных, которые традиционно считались сферой применения качественных методов филологического и коммуникативного анализа: художественные тексты и устная быденная речь.

2. Для получения новых знаний об изменениях, происходящих в социально-лингвокультурном пространстве современного русского языка: в социо-речевых сферах устной разговорной речи и художественных текстов (профессиональных и “наивных” авторов) формируются и манифестируются важнейшие паттерны мировоззренческого и быденного осмысления социальной реальности, проявляются языковые и поведенческие сдвиги, в том числе — конвергентные процессы, происходящие между разными жанрами, сферами, регистрами общения.

3. Для создания научно-прикладных разработок на основе созданной ресурсной базы: сервисов для автоматического анализа и синтеза русской речи во всех сферах коммуникации, а также разработки интеллектуальных систем для прогнозирования коммуникативных рисков и неудач.

Совокупность данных факторов позволит создать на базе подразделения конкурентоспособный научно-исследовательский центр.



#### 4. Цель создания.

Формирование исследовательской инфраструктуры для проведения прикладных лингвистических и филологических исследований, с одной стороны, связанных с моделированием конвергентных процессов, происходящие между разными жанрами, сферами, регистрами общения, а с другой — направленных на апробацию современных компьютерных методов обработки больших данных, включая методы машинного и глубокого обучения, к большим объемам текстовой и звуковой информации, а также методов искусственного порождения текстов — как письменных, так и устных, — и выявлении оптимальных подходов работы с большими текстовыми данными указанных жанров.

#### 5. Задачи.

- Создание условий для проведения лингвистических и филологических исследований в интересах научных структурных подразделений НИУ ВШЭ, а также сторонних заказчиков.
- Организационное и методическое сопровождение прикладных лингвистических исследований в области обработки и анализа больших текстовых данных и звучащей речи.
- Разработка новых электронных лингвистических и филологических ресурсов (письменных и устных) для решения актуальных теоретических и практических задач как традиционной, так и компьютерной лингвистики.
- Апробация современных компьютерных методов обработки больших данных, включая методы машинного и глубокого обучения, к большим объемам текстовой и звуковой информации.
- Решение фундаментальных филологических задач, связанных с моделированием нарративной реальности письменных и устных текстов разных социо-речевых сфер, выявление их стилистических особенностей и степени их вариативности.
- Моделирование конвергентных процессов, наблюдаемых на разных лингвистических уровнях между текстами разных жанров и разных социо-речевых сфер.
- Апробация методов искусственного порождения текстов — как письменных, так и устных, — обученных на больших текстовых данных, выявление их «слабых мест».
- Подготовка практических рекомендаций для оптимизации систем искусственного интеллекта, связанных с обработкой и генерации текстов/звучащей речи.

#### 6. Описание научно-исследовательской деятельности.

Лаборатория языковой конвергенции факультета Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ – Санкт-Петербург будет:

- формировать коллекции текстовых и мультимедийных данных, отражающих современные конвергентные процессы, происходящие между различными социо-речевыми сферами использования русского языка (корпусы, датасеты);
- проводить анализ текстовых коллекций и корпусов звучащей речи с применением различных методов NLP;
- генерировать новое знание о конвергентных процессах в разнообразных социо-речевых сферах использования языка;
- использовать данное знание для разработки продуктов в области NLP и NLU.

#### **7. Описание иной деятельности.**

Студенты НИУ ВШЭ – Санкт-Петербург смогут проходить учебную и производственную практику в Лаборатории языковой конвергенции СПбШГНИИ, участвовать в проектной деятельности на базе лаборатории.

#### **8. Взаимодействие со структурными подразделениями НИУ ВШЭ, а также сторонними организациями.**

Лаборатория языковой конвергенции факультета Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург будет осуществлять взаимодействие по реализации фундаментальных и прикладных проектов внутри НИУ ВШЭ и за его пределами, в частности для решения таких задач, как сбор и обработка больших текстовых данных, обработка больших массивов речевых данных, создание новых лингвистических и филологических ресурсов, проведение совместных исследований и научных мероприятий, участие в чтении учебных курсов и образовательных лекций, реализация совместных проектов и практики для бакалавров и магистров в кампусах Москвы и Санкт-Петербурга, привлечение студентов к участию в лингвистических экспедициях, разработка популярного образовательного контента по основным направлениям деятельности научного партнера.

#### **9. Ожидаемые результаты от создания.**

Основными результатами создания лаборатории языковой конвергенции факультета Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург будут:

- Разработанные в рамках проекта электронные лингвистические и филологические ресурсы (корпус малой русской прозы (русского рассказа) 1930-2000 гг.; устный корпус современной повседневной речи молодежи).
- Результаты апробации современных компьютерных методов обработки больших данных, включая методы машинного и глубокого обучения, к большим объемам лингвистической/филологической информации – как текстовой, так и звуковой, оценка применимости этих методов к лингвистическим/филологическим данным разных жанров и разных социо-речевых сфер.
- Научное описание русского художественного текста малой формы (рассказа) как модели нарративного повествования, выполненное на материале художественных текстов XX-XXI гг., написанных как профессиональными писателями, так и непрофессиональными авторами.



- Научное описание особенностей современной повседневной речевой коммуникации, выполненное на материале звукозаписей речи и/или общения в социальных сетях.
- Описание моделей конвергентных процессов, наблюдаемых на разных лингвистических уровнях между текстами разных жанров и разных социо-речевых сфер, построенных с использованием компьютерных методов.
- Результаты апробации методов искусственного порождения текстов — как письменных, так и устных, — обученных на больших текстовых данных.
- Практические рекомендации для оптимизации систем искусственного интеллекта, связанных с обработкой и генерацией текстов/звучащей речи рассматриваемых жанров.

#### 10. Финансирование деятельности.

Финансирование деятельности лаборатории языковой конвергенции будет осуществляться за счет следующих источников:

- средств, выделенных факультету Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург на выполнение работ в рамках проекта «Текст как Big Data: моделирование конвергентных процессов в языке и речи цифровыми методами», поддержанного по результатам конкурса проектов на выполнение фундаментальных научных исследований распределенными межкампусными научными подразделениями Национального исследовательского университета «Высшая школа экономики» на основании служебных записок от руководителя проекта;
- средства, получаемые от приносящей доход деятельности (научно-исследовательской, экспертно-аналитической).

#### 11. Сведения о кадровом составе.

Кадровый состав.

| Наименование должностей  | Количество штатных единиц |
|--------------------------|---------------------------|
| Руководитель лаборатории | 0,1                       |
| Зам.руководителя         | 0,1                       |
| Стажер-исследователь     | 1                         |
| Итого                    | 1,2                       |

#### 12. Потребности в площадях, оборудовании.

Планируется, что лаборатория языковой конвергенции факультета Санкт-Петербургская школа гуманитарных наук и искусств НИУ ВШЭ - Санкт-Петербург будет располагаться на факультете Санкт-Петербургская школа гуманитарных наук и искусств по адресу: Санкт-Петербург, наб. канала Грибоедова, 123. Для работы лаборатории желательно наличие помещения с 5 рабочими местами, оборудованными ПК, МФУ.

## Концепция

научной лаборатории Интернета вещей и киберфизических систем факультета Санкт-Петербургская школа физико-математических и компьютерных наук НИУ ВШЭ - Санкт-Петербург

### 1. Предпосылки создания.

Отсутствие собственного подразделения, способного вести исследования в области Интернета вещей ограничивает возможности научных коллективов НИУ ВШЭ СПб участвовать в конкурсах на проведение полномасштабных социологических исследований.

### 2. Состояние предметной области в России и за рубежом.

Специфические особенности систем Интернета вещей в сценарии массовой межмашинной связи (ограничение на энергопотребление, короткое сообщение и др.) не позволяет в полной мере использовать результаты существующих исследований в области случайного доступа, помехоустойчивого кодирования. Это делает актуальной разработку новых моделей и алгоритмов, учитывающих специфику таких систем, анализ их стабильности, защищённости от влияния естественных и искусственных помех и вычислительной сложности используемых методов.

### 3. Обоснование создания.

Факультету Санкт-Петербургская школа физико-математических и компьютерных наук НИУ ВШЭ - Санкт-Петербург необходима собственная структурная единица для организации исследований в области Интернета вещей.

### 4. Цель создания.

Формирование исследовательской инфраструктуры для разработки способов повышения энергетической эффективности в беспроводных сетях с низким энергопотреблением.

### 5. Задачи.

- Провести обзор технологий передачи данных в беспроводных сетях с низким энергопотреблением и выбрать конкретную технологию как объект для дальнейшего исследования.
- Предложить способ повышения вероятности доставки сообщений с учётом специфических особенностей выбранной технологии.
- Разработать методику проведения натуральных экспериментов для проверки работоспособности предложенных способов повышения вероятности доставки сообщений.
- Рассмотреть вопросы обеспечения энергетической и экономической эффективности предложенных способов повышения вероятности доставки сообщений.



#### **6. Описание научно-исследовательской деятельности.**

Исследование влияния ограничения затрат энергии на характеристики (условия стабильности, помехозащищённости) системы при большом количестве абонентских устройств. Учёт данного ограничения является новизной проекта по сравнению с другими исследованиями условий стабильности систем случайного множественного доступа.

#### **7. Описание иной деятельности.**

Студенты НИУ ВШЭ смогут проходить производственную практику в научной лаборатории.

#### **8. Взаимодействие со структурными подразделениями НИУ ВШЭ, а также сторонними организациями.**

Научная лаборатория Интернета вещей и киберфизических систем будет осуществлять взаимодействие по реализации фундаментальных и прикладных проектов внутри НИУ ВШЭ и за его пределами.

#### **9. Ожидаемые результаты от создания.**

Основными результатами являются новые модели систем и алгоритмы случайного множественного доступа, помехоустойчивого кодирования и декодирования, защиты информации, теоретически обоснованные границы достижимости для затрачиваемой энергии на передачу одного бита информации в стабильной системе.

Полученные результаты позволят оценить затраты энергии при использовании различных беспроводных технологий в системах Интернета вещей для сценария массовой межмашинной связи и могут быть использованы при разработке систем данного типа.

#### **10. Финансирование деятельности.**

Финансирование деятельности научной лаборатории Интернета вещей и киберфизических систем будет осуществляться за счет следующих источников:

- средств, выделенных НИУ ВШЭ на выполнение работ в рамках научного проекта «Исследование и разработка методов множественного доступа и помехоустойчивого кодирования для организации энергоэффективной передачи данных в системах Интернета вещей», поддержанного по результатам конкурса проектов на выполнение фундаментальных научных исследований распределёнными межкампусными научными подразделениями НИУ ВШЭ;
- средства, получаемые от приносящей доход деятельности (научно-исследовательской и образовательной).

#### **11. Сведения о кадровом составе.**

Кадровый состав.

| Наименование должностей   | Количество штатных единиц |
|---------------------------|---------------------------|
| Заведующий лабораторией   | 0,1                       |
| Младший научный сотрудник | 0,3                       |
| Стажер-исследователь      | 0,3                       |
| Итого                     | 0,7                       |

## 12. Потребности в площадях, оборудовании.

1. Помещение, оборудованное мебелью из расчёта 10 рабочих мест.
2. Оборудование: интерактивная доска; МФУ; 2 графических планшета (WACOM); 2 ноутбука; ПК на 10 рабочих мест.