

# Александра Шикунова

Москва, Россия  
Год рождения: 2002

Контакты:

[Telegram](#)

[Email](#)

[Github](#)

[Сайт](#)

[В Вышке](#)

## Навыки

**Языки:** Python, SQL, R

**Библиотеки/фреймворки:** Huggingface, Pytorch, sklearn, pandas, numpy, nltk, PySpark, FastAPI, Flask, streamlit

## Опыт работы

- с 2023 **NLP-инженер, ПАО Сбербанк, центр ИИ блока “Сервисы”**
- ✳ обучила SFT SLM с 400М параметров суммаризировать новости в 17 раз быстрее 7B модели при сравнимом качестве
  - ✳ построила единый пайплайн для задач инфопоиска, тестирующий несколько подходов (tf-idf, BM25, эмбеддеры)
  - ✳ для ускорения обработки новостного потока дедублировала, кластеризовала и суммаризовала новости
  - ✳ сводила задачи NER и поиска по реестру товаров к regex, над продуктивизацией поиска работала со стажером
- 2022-2023 **Аналитик больших данных, ПАО Мегафон, команда чат-бота**
- ✳ регулярное обновление продовых моделей intent & entity recognition
  - ✳ обернула модель кластеризации для разведки новых интенгов в приложение для ассессоров
- с 2021 **Стажер-исследователь в НУЛ по формальным моделям в лингвистике, ВШЭ**
- 2020-2023 **Преподавание олимпиадного английского в образовательном центре “Взлёт”**
- 2023 **Учебный ассистент по теории вероятности и математической статистике, Школа Лингвистики ВШЭ**
- 2022 **Учебный ассистент по Python на майноре “Интеллектуальный анализ данных”, ФКН ВШЭ**
- 2021 **Учебный ассистент по дискретной математике, кафедра высшей математики ВШЭ**

## Образование

- 2024-2026 **Магистратура “Лингвистическая теория и описание языка”, ФГН НИУ ВШЭ, Москва**
- 2020-2024 **Бакалаврская программа “Фундаментальная и компьютерная лингвистика”, ФГН НИУ ВШЭ, Москва**
- 2021-2023 **Майнор “Интеллектуальный анализ данных”, ФКН НИУ ВШЭ**

## Личные проекты

- с 2021 **Полевые исследования малых языков России: акустическая фонетика, морфосинтаксис**
- 2024 **Обработка данных и построение инфографик для изучения фонологии [мокшанского](#) и [лесного ненецкого](#)**
- 2023-2024 **[NonceRNN](#) – генерация псевдослов с помощью RNN**
- 2022 **[MessageHDD](#) – метрика лексического разнообразия для корпуса маленьких текстов**
- 2021 **[Анализ данных синтаксического эксперимента на Python](#)**

## Языки

- ✳ Русский родной
- ✳ Немецкий, персидский на начальном уровне
- ✳ Английский TOEFL C2 (2023)
- ✳ Мокшанский, севернохантыйский, лесной
- ✳ Французский средний уровень
- ✳ ненецкий полевая работа