

ЭЛЬМИРА МУСТАКИМОВА

Дата рождения: 19.01.1995

Москва, Россия

Моб.телефон:

+7 (926) 908 06 56

Email:

egmustakimova_2@edu.hse.ru

Веб:

www.hse.ru/staff/egmustakimova

Прочитайте это резюме онлайн на www.elmiram.github.io/resume_rus.html

ОБРАЗОВАНИЕ

БАКАЛАВРИАТ

2012 – настоящее время

Национальный Исследовательский Университет «Высшая Школа Экономики» (НИУ ВШЭ), Москва, Россия

В настоящее время – студентка выпускного курса

Образовательная программа: Фундаментальная и компьютерная лингвистика

Основные курсы:

- Теория языка (фонетика, морфология, семантика, синтаксис, дискурс)
- Экспериментальная лингвистика и нейролингвистика
- Языковое разнообразие и типология
- Программирование (Python, R)
- Алгоритмы и структуры данных
- Базы данных
- Автоматическая обработка естественного языка

ЛЕТНИЕ ШКОЛЫ

Июль 2014 – Летняя школа языка идиш "Yiddish on the Continent"

Рабочий язык: французский, идиш

Свободный университет Брюсселя, Брюссель, Бельгия

Сентябрь 2014 – Летняя школа "Введение в современную нейролингвистику"

Рабочий язык: английский

НИУ ВШЭ, Москва, Россия

Июнь 2015 – Летняя школа "Местоимения: синтаксис, семантика, обработка"

Рабочий язык: английский

НИУ ВШЭ, Москва, Россия

ОПЫТ РАБОТЫ

НОЯБРЬ 2013 – НАСТОЯЩЕЕ ВРЕМЯ УЧЕБНЫЙ АССИСТЕНТ

Школа лингвистики, Национальный Исследовательский Университет “Высшая Школа Экономики”

Курсы: *Программирование для лингвистов* в 2013, *Программирование (Python)* с 2014 по настоящее время

Обязанности:

- Проверка и оценивание домашних заданий и контрольных работ по программированию
- Проведение консультаций для небольших групп студентов
- Регулярные отчеты преподавателю об успехах и основных ошибках студентов

ФЕВРАЛЬ 2014 – НАСТОЯЩЕЕ ВРЕМЯ

ЛАБОРАНТ

Лингвистическая лаборатория по корпусным технологиям, Национальный Исследовательский Университет “Высшая Школа Экономики”

Обязанности:

- Разработка современной платформы для двух языковых ресурсов (корпусов): создание поискового алгоритма и системы аннотации документов
- Работа с Python, Django и MySQL
- Обновление веб-страниц разрабатываемых ресурсов
- Оказание технической поддержки пользователям корпусов

НАВЫКИ

Программирование

- Языки: Python, R, JavaScript
- Фреймворк: Django, Flask, D3.js
- Базы данных: SQL (MySQL, SQLite), NoSQL (ElasticSearch, Neo4j)
- IDE: PyCharm
- LaTeX
- Git

Языки

- Русский – родной
- Английский – продвинутый (IELTS 115/120)
- Немецкий, французский, идиш – средний
- Японский, индонезийский – начинающий

ПРОЕКТЫ

2014 – НАСТОЯЩЕЕ ВРЕМЯ

«ИССЛЕДОВАНИЯ В ЛЕКСИЧЕСКОЙ ТИПОЛОГИИ»

Исследование глаголов падения в языках мира. В этом проекте я занимаюсь сбором и анализом данных 4 языков (индонезийский, идиш, татарский, японский). Кроме этого, я разрабатываю веб-интерфейс для визуализации базы данных, который будет доступен по ссылке www.web-corpora.net/falling.

2014 – НАСТОЯЩЕЕ ВРЕМЯ УЧЕБНЫЙ КОРПУС»

«КОРПУС РУССКИХ СТУДЕНЧЕСКИХ ТЕКСТОВ» И «РУССКИЙ

Корпус русских студенческих текстов (www.web-corpora.net/learner_corpus) содержит тексты, написанные студентами, носителями русского языка, в академической среде. Русский учебный корпус (www.web-corpora.net/heritage_corpus) содержит тексты, написанные изучающими русский язык студентами с разными родными языками (эритажные носители и носители других языков). Я разрабатываю онлайн платформу для этих двух корпусов. Моя цель – создать сложную систему с возможностью добавления новых текстов, аннотаций, создания подкорпусов и поиска по ним.

НОЯБРЬ 2014 – МАРТ 2015 СТУДЕНЧЕСКИЙ СЕМИНАР ПО ЛИНГВИСТИКЕ

В какой-то момент я поняла, что мало знаю об исследованиях, в которых участвуют мои однокурсники и другие студенты. Поэтому я предложила собирать ежемесячный студенческий семинар, на котором каждый мог бы рассказать о своем исследовании и получить комментарии. Довольно много студентов поддержали эту идею, и таким образом я стала председателем студенческих семинаров по лингвистике.

2014 «АВТОМАТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ СЕМАНТИЧЕСКОГО ОРЕОЛА МЕТРА»

“Автоматическое определение семантического ореола метра” это итоговый проект на курсе “Метрика в русской поэзии”. Эта работа вдохновлена идеей о том, что конкретный поэтический метр в сознании поэта связан с определенным кругом тем. Эта идея была довольно популярна среди некоторых русских исследователей XX-го века. Я попробовала использовать различные методы машинного обучения для кластеризации русских стихотворений по темам, чтобы увидеть, есть ли корреляция тематических кластеров с группами стихотворений одного метра. Хотя ответ на этот вопрос оказался отрицательным, работа привлекла внимание филологов, и я выступила с докладом об этом эксперименте на международной конференции молодых филологов в Тарту.

2014 «АВТОМАТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ДЕФЕКТНЫХ ПАРАДИГМ НА ОСНОВЕ КОРПУСОВ»

“Автоматическое определение дефектных парадигм на основе корпусов” – это тема моей курсовой работы третьего курса. В то время меня интересовали статьи Адама Албрайта о дефектных парадигмах, и я тогда же заинтересовалась машинным обучением в лингвистических исследованиях. Эти два интереса привели к проекту, в котором я использовала деревья решений, классификатор KNN и наивный байесовский классификатор, чтобы автоматически определить на основе корпусных данных, является ли конкретное слово дефектным или нет.

2013 – 2014 «СОЗДАНИЕ КОРПУСНЫХ ИНСТРУМЕНТОВ ДЛЯ ЯЗЫКА ИДИШ»

Целью проекта являлось создание инструментов для работы с языком идиш, таких как транслитератор, автоматическое выравнивание текстов и аудиофайлов на идише, создание аннотированного корпуса. Моя задача в проекте состояла в автоматическом снятии морфологической омонимии в текстах на идише. Дизамбигуатор написан на языке Python. Я использовала скрытые марковские модели, алгоритм Брилла и вручную написанные правила.

2013 – 2014 «БАЗА ДАННЫХ КАК ИНСТРУМЕНТ ЛЕКСИКО-ТИПОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ: ГЛАГОЛЫ ЗВУКОВ ЖИВОТНЫХ И МОДЕЛИ ИХ СЕМАНТИЧЕСКОЙ ДЕРИВАЦИИ»

Цель проекта заключалась в создании базы данных глаголов звуков животных и их метафор. В этом проекте я занималась сбором данных для двух языков, а также разработкой веб-интерфейса для визуализации базы данных (www.web-corpora.net/zvukimu).

ПУБЛИКАЦИИ

Kuzmenko, Elizaveta, Mustakimova, Elmira. Syntactic Phraseme of the Type X-t' ne pereX-t' // In [Proceedings of the 6th International Conference on Meaning-Text Theory](#)

Kuzmenko E., Mustakimova E. [Automatic Disambiguation in the Corpora of Modern Greek and Yiddish](#), in: Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции "Диалог" (2015) / Отв. ред.: В. Селегей. [б.и.], 2015. P. 388-398.

Мустакимова Э. Г., Орехов Б. В. Глаголы звуков животных в башкирском и татарском языках // В кн.: «И хрюкотали зелюки: глаголы звуков животных - типология метафор» / Отв. ред.: Т. И. Резникова, А.С. Выренкова, Б. В. Орехов, Е. В. Рахилина. Языки славянских культур, 2015. (в печати)

Мустакимова Э. Г. Глаголы звуков животных в индонезийском языке // В кн.: «И хрюкотали зелюки: глаголы звуков животных - типология метафор» / Отв. ред.: Т. И. Резникова, А.С. Выренкова, Б. В. Орехов, Е. В. Рахилина. Языки славянских культур, 2015. Гл. 16. (в печати)

КОНФЕРЕНЦИИ

2015

Октябрь 14 – 16. 3-я конференция по общему, скандинавскому и славянскому языкознанию, GenSLing 2015 (Москва). Доклад: *Глаголы падения в типологической перспективе*

Октябрь 10 – 11. ConCorT Junior. Корпусные технологии и компьютерные методы в современной гуманитарной науке (Москва). Доклад: *Глаголы звуков животных: визуализация базы данных*

Июль 20 – 24. Corpus Linguistics 2015 (Ланкастер, Великобритания). Доклад: *Corpus of Russian Student Texts: goals, annotation, and perspectives*

Июнь 16 – 19. Pronouns: Syntax, Semantics, Processing (Москва). Доклад: *Automatic Identification of Paradigm Gaps in Corpora*

Май 27 – 30. Международная конференция "Диалог 2015": Компьютерная лингвистика и интеллектуальные технологии (Москва). Доклад: *Automatic Disambiguation in the Corpora of Modern Greek and Yiddish*

Апрель 24 – 27. Международная конференция молодых филологов в Тарту (Тарту). Доклад: *Автоматическое определение семантического ореола метра*

Апрель 14. Студенческая конференция Школы Лингвистики НИУ ВШЭ и ИИЯ МГПУ (Москва). Доклады: *Корпус русских учебных текстов* и *Автоматическое определение дефектных парадигм по данным корпусов*

2014

Декабрь 1 – 3. Семинар по когнитивной семантике и лексической типологии (Хельсинки). Доклад: *Анонс нового типологического проекта: типология глаголов падения*

Сентябрь 26 – 27. Бестиарный код культуры (RES ET VERBA - 4) (Москва). Доклад: *Грустное рычание и бодрое мычание: какие звуки издают животные в разных языках*

ХОББИ

- Рукоделие, в основном из бумаги. Занимаюсь разными видами рукоделия из бумаги: оригами, аппликация, создание открыток, скрапбукинг.
- Создание книг (Bookbinding). Под этим я имею в виду довольно долгий процесс сшивания листов бумаги, приклеивания форзацев, проектирование и сбор обложки из картона и тканей. Обычно я делаю пустые книги-блокноты, которые могу потом преподнести кому-то в подарок, но иногда я распечатаваю тексты и собираю настоящую книгу. Я люблю ощущение, которое возникает, когда книга готова: из ничего ты создал нечто красивое и полезное.
- Катание на коньках и роликах.
- Изучение японского языка.