

МИНОБРНАУКИ РОССИИ
Федеральное государственное автономное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"

Авторы-составители: **Исаева Екатерина Владимировна**

Программа учебной практики
SMART RESEARCH IN LINGUISTICS
Код УМК 100695

Утверждено
Протокол №9
от «12» мая 2023 г.

Пермь, 2023

1. Вид практики, способ и форма проведения практики

Вид практики **учебная**

Тип практики **учебно-ознакомительная практика**

Способ проведения практики **стационарная, выездная**

Форма (формы) проведения практики **дискретная**

2. Место практики в структуре образовательной программы

Учебная практика « Smart Research in Linguistics » входит в вариативную часть Блока « М.2 » образовательной программы по направлениям подготовки (специальностям):

Направление: **45.04.02** Лингвистика

направленность Цифровая лингвистика и перевод

Цель практики :

Develop skills for automation of linguistic research applied to the students own projects.

The practical course should be taken after a successful completion of the Computer Linguistics for International Professional Communication course.

Задачи практики :

1. Master the instruments for NLP analysis;
2. To practice formulating research questions for a smart linguistic research;
3. To practice designing a smart linguistic research;
4. To carry out a smart linguistic research;
5. To present the results of the study in a form of a research paper;
6. To present the results of the study in a form of an oral presentation.

3. Перечень планируемых результатов обучения

В результате прохождения практики **Smart Research in Linguistics** у обучающегося должны быть сформированы следующие компетенции:

45.04.02 Лингвистика (направленность : Цифровая лингвистика и перевод)

ОПК.6 Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию

Индикаторы

ОПК.6.1 Составляет и оформляет научную документацию (диссертация, отчеты, обзоры, аннотации, доклады, статьи, рефераты), библиографии и ссылки в соответствии с принятыми нормами и стандартами

ОПК.7 Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации

Индикаторы

ОПК.7.1 Использует информационно-поисковые системы открытого доступа для проведения экспертного анализа дискурса, средства автоматизированной обработки вербальной информации

ПК.3 Способен проводить самостоятельные научные исследования по актуальной проблематике в области лингвистики, переводоведения и межкультурной коммуникации

Индикаторы

ПК.3.2 Применяет современные методики поиска, анализа и обработки материала исследования на основе системного представления о динамике развития переводоведения, лингвистики и межкультурной коммуникации

4. Содержание и объем практики, формы отчетности

This practical course is designed to equip students with the necessary skills and knowledge to conduct effective research in the field of linguistics using smart tools for natural language processing (NLP). Through a combination of tutorials and homework assignments, students will refresh their knowledge of NLP systems, practice in formulating research questions, designing studies, collecting and analyzing data automatically, making data-driven decision, and presenting their findings in a clear and concise manner.

Направление подготовки	45.04.02 Лингвистика (направленность: Цифровая лингвистика и перевод)
форма обучения	очная
№№ триместров, выделенных для прохождения практики	2
Объем практики (з.е.)	3
Объем практики (ак.час.)	108
Форма отчетности	Экзамен (2 триместр)

Примерный график прохождения практики

Количество часов	Содержание работ	Место проведения
36	<p>1. NLP instruments for linguists</p> <p>NLP Instruments for Linguists in the Course 'Smart Research in Linguistics'</p> <p>In the practical course "Smart Research in Linguistics," the section on "NLP Instruments for Linguists" focuses on introducing students to essential tools and platforms used in Natural Language Processing (NLP) that are beneficial for linguistic research. By incorporating these tools, linguists can enhance their research capabilities and efficiency. Some key points covered in this section include:</p> <ul style="list-style-type: none"> • Overview of Tools and Platforms: The course introduces students to a range of NLP tools such as Gensim, SpaCy, IBM Watson, Natural Language Toolkit (NLTK), MonkeyLearn, TextBlob, and Stanford CoreNLP 1, 5 These tools offer functionalities like text analysis, sentiment analysis, keyword extraction, and more. • Automation Capabilities: NLP tools enable automation in data analysis, allowing for quick and thorough analysis of text and speech data 1 This automation aids linguists in processing large volumes of data efficiently. • Applications of NLP: The section highlights how NLP can help linguists resolve language ambiguities, provide structure to data, aid in textual analytics, speech recognition, and even investigative discovery by spotting patterns in written reports or emails. • Keyword Extraction and Sentiment Analysis: Students will learn 	Department of Professional English Language

Количество часов	Содержание работ	Место проведения
	<p>about the importance of keyword extraction algorithms to summarize text and sentiment analysis techniques to categorize text into positive, negative, or neutral sentiments. By familiarizing students with these NLP instruments, the course equips them with practical skills to leverage technology for linguistic research, enhancing their ability to analyze language data effectively.</p>	
2. Smart research design		
36	<p>This section delves into the strategic planning and methodology involved in conducting research in linguistics effectively. Key components of this section include: analysing sample NLP workflows in Knime and Orange or notebooks in Google Colab; customizing the workflow or the code for current tasks; carrying out the research with their own data.</p> <p>Orange and KNIME are valuable tools that assist linguists in Natural Language Processing (NLP) by providing functionalities for text analysis, data processing, and visualization. Here's how these tools help linguists in NLP:</p> <p>-Orange:</p> <p>Orange is a data visualization and analysis tool that offers a user-friendly interface for data mining, machine learning, and NLP tasks. Linguists can utilize Orange for tasks like text analysis, sentiment analysis, topic modeling, and more, making it a versatile tool for linguistic research.</p> <p>- KNIME:</p> <p>KNIME is a platform that supports NLP tasks and offers language packages for text processing.</p> <p>While KNIME's enrichment node primarily supports English, users can implement tagging, cleaning, stemming, lemmatizing in non-English languages like Japanese and Vietnamese using Python and integrate these processes back into KNIME for further analysis.</p> <p>KNIME also provides extensions like spaCy for NLP tasks, allowing users to transition from Python to KNIME seamlessly for NLP projects.</p> <p>By leveraging the capabilities of Orange and KNIME, linguists can streamline their NLP workflows, analyze linguistic data effectively, and extract valuable insights from text data in various languages. These tools enhance the efficiency and accuracy of linguistic research processes, enabling researchers to delve deeper into the complexities of language processing.</p>	Department of Professional English Language
3. Writing and Presenting Research		
36	<p>This section emphasizes the importance of planning and organizing research content before writing, including creating a research background, drafting tables and figures, and committing</p>	Department of Professional English Language

Количество часов	Содержание работ	Место проведения
	<p>ideas to paper. Students will also learn how to present their research orally.</p> <p>The section starts with brief introduction into the IMRAD structure, add-ons for guided writing, and bibliographic managers.</p> <p>The practical part of the section is devoted to writing a report and presenting on the smart research findings.</p> <p>By incorporating these elements into the course curriculum, students will develop the necessary skills to effectively communicate their linguistic research findings to various audiences, both in written and oral formats. This knowledge will enable them to disseminate their research in a clear, concise, and impactful manner, enhancing their ability to contribute to the field of linguistics.</p>	

5. Перечень учебной литературы, необходимой для проведения практики

Основная

1. Лысаков, К. Ф. Практическое программирование на Python : учебное пособие / К. Ф. Лысаков. — Новосибирск : Новосибирский государственный университет, 2023. — 75 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/134584.html>
2. Титов, А. Н. Визуализация данных в Python. Работа с библиотекой Seaborn : учебно-методическое пособие / А. Н. Титов, Р. Ф. Тагиева. — Казань : Издательство КНИТУ, 2023. — 144 с. — ISBN 978-5-7882-3326-0. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/136145.html>

Дополнительная

1. Титов, А. Н. Python. Обработка данных : учебно-методическое пособие / А. Н. Титов, Р. Ф. Тагиева. — Казань : Издательство КНИТУ, 2022. — 104 с. — ISBN 978-5-7882-3171-6. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <http://www.iprbookshop.ru/129220.html>
2. Терминологический инжиниринг на иностранном языке. Осваиваем терминологию компьютерной безопасности: учебное пособие для студентов, обучающихся по направлению подготовки магистров «Фундаментальная и прикладная лингвистика»/М-во науки и высш. образования РФ, Перм. гос. нац. исслед. ун-т.-Пермь:ПГНИУ,2020, ISBN 978-5-7944-3429-3.-145. <https://elis.psu.ru/node/595731>

6. Перечень ресурсов сети «Интернет», требуемых для проведения практики

При прохождении практики требуется использование следующих ресурсов сети «Интернет» :

<https://www.nobledesktop.com/classes-near-me/blog/best-natural-language-processing-tools> 8 Best Tools for Natural Language Processing in 2024

https://studies.helsinki.fi/courses/course-implementation/otm-d3a1c8dd-795f-4557-838b-6636c256e13a/ENG-Ling353/Natural_Language_Processing_for_Linguists Natural Language Processing for Linguists, Open Uni: Natural Language Processing for Linguists 5 cr

<https://www.kdnuggets.com/2022/11/5-linguistics-courses-nlp-practitioners.html> 5 Linguistics Courses for NLP Practitioners

https://www.digitalhumanities.unito.it/do/corsi.pl/Show?_id=nxwv Linguistic Resources for Natural Language Processing

<https://aimagazine.com/top10/top-10-natural-language-processing-tools-and-platforms> Top 10 Natural Language Processing tools and platforms

<https://www.knime.com/events/from-python-to-knime-spacy> From Python to Knime SpaCy

<https://hub.knime.com/search?q=language%20pack&sort=best&tag=natural%20language%20processing&type=All> Knime Hub

<https://blogs.helsinki.fi/language-technology/> Language Technology in Helsinki

7. Перечень информационных технологий, используемых при проведении практики

Образовательный процесс по практике **Smart Research in Linguistics** предполагает использование следующего программного обеспечения и информационных справочных систем:

The educational process for the course involves the use of the following software and information systems:

- presentation aids (slides on the topics of lectures and practical classes);
- on-line access to the Electronic Library System (ELS);
- access to the electronic information and educational environment of the university.

List of necessary licensed and (or) freely distributed software:

- software for data analysis: Knime, Orange;
- programming languages and environment: Google Colab, Python 3.10, 3.11, Anaconda
- application that allows viewing and playing media content PDF files "Adobe Acrobat Reader DC";
- software, demonstrating video materials (player) "WindowsMediaPlayer";
- program for viewing Internet content (Google Chrome, Yandex Browser)
- Office suite "LibreOffice" or Microsoft Office 2003 Professional.

Reference-legal program ConsultantPlus Additionally can be used:

- Sanako Lab 250 licensed software;
- Abby Lingvo electronic dictionary.

Open access software: Hot Potatoes, Quizlet, Trello, XMind, Miro, Figma, Orange.

When mastering the material and performing assignments on the discipline, it is recommended to use the materials placed in the personal student accounts in the ETIS PSU (student.psu.ru).

8. Описание материально-технической базы, необходимой для проведения практики

When organizing remote work and conducting classes online can be used:

The material and technical base is assured by the following:

- 1) video conferencing system based on BigBlueButton platform (<https://bigbluebutton.org/>);
- 1) Laboratory classes are held in a classroom equipped with presentation equipment (projector, screen, computer/laptop) with appropriate software and (or) computer class (hardware and software equipment is audio and video content, as well as tests, checked tasks, tasks for joint work. defined in the passport of the computer class).
- 2) testing system Indigo (<https://indigotech.ru/>).
- 2) For independent work - a classroom equipped with computers with the ability to connect to the "Internet", provided with access to the electronic information and educational environment of the university. Facilities of

the Scientific Library of PSU.

- 3) При освещении carried out in a classroom equipped with presentation equipment (проектор, экран, компьютер, Маркер) with appropriate software обучаю (о) щая / ПГНИУ (student.psu.ru). Faculty of the PSU research library provide students with online режим online могут be used in the Bibliographical Department, Bld.1, room 142. Equipped with 3 personal computers with access to local and global computer networks на основе платформы BigBlueButton (<https://bigbluebutton.org/>).
2. Humanitarian Reading Room (Bld. 2, room 418) equipped with computers with access to local and global computer networks презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания reading hall of natural science literature, Bld. 6, room 107a. Equipped with 5 personal computers with access to local and global Indigit (<https://onki.gotech.ru/>).
4. Department of Foreign Literature, Bld.2, room 207. One personal computer with access to local and global computer network.
5. Library of the Faculty of Law, Bld.9, room 4. Equipped with 11 personal computers with access to local and global computer networks.
6. Reading room of the Faculty of Geography, Bld.8, room 419. Equipped with 6 personal computers with access to local and global computer networks.

All computers installed in the scientific library are equipped with the following software:

ALT Linux operating system; Libreoffice office package.

Reference legal system "ConsultantPlus"

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютера с доступом к локальной и глобальной компьютерным сетям.
2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.
3. Читальный зал естественной литературы, корп.6, ауд. 107a. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.
4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с доступом к локальной и глобальной компьютерным сетям.
5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.
6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

9. Методические указания для обучающихся по прохождению практики

Фонды оценочных средств для проведения промежуточной аттестации

Планируемые результаты обучения по практике для формирования компетенции. Индикаторы и критерии их оценивания

ОПК.7

Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.7.1 Использует информационно-поисковые системы открытого доступа для проведения экспертного анализа дискурса, средства автоматизированной обработки вербальной информации</p>	<p>Ability to use open access information retrieval systems for expert discourse analysis and means of automated processing of verbal information.</p>	<p style="text-align: center;">Неудовлетворительно</p> <p>The required skills have not been developed.</p> <p style="text-align: center;">Удовлетворительно</p> <p>The student knows at least two modern open access information retrieval systems or means of automated processing of verbal information. The student cannot analyse discourse using modern information retrieval systems. The student can update and apply sampled workflows for simple NLP tasks.</p> <p style="text-align: center;">Хорошо</p> <p>The student knows modern open access information retrieval systems and means of automated processing of verbal information. The student cannot analyse discourse using modern information retrieval systems. The student applies means of automated processing for NLP tasks.</p> <p style="text-align: center;">Отлично</p> <p>The student knows modern open access information retrieval systems and means of automated processing of verbal information. The student can analyse discourse using modern information retrieval systems. The student applies means of automated processing for NLP tasks.</p>

ОПК.6

Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.6.1 Составляет и оформляет научную документацию</p>	<p>Ability to use the results of smart research to justify and illustrate scientific papers, to present the results of the study in accordance</p>	<p style="text-align: center;">Неудовлетворительно</p> <p>The required skills have not been developed.</p> <p style="text-align: center;">Удовлетворительно</p> <p>The student does not know the requirements for</p>

<p>(диссертация, отчеты, обзоры, аннотации, доклады, статьи, рефераты), библиографии и ссылки в соответствии с принятыми нормами и стандартами</p>	<p>with international requirements.</p>	<p style="text-align: center;">Удовлетворительно</p> <p>the international presentation of the research data in a written form or is not able to use bibliographic managers to provide references in various formats.</p> <p>The student can use the results of smart research to justify and illustrate scientific papers but cannot meet the international requirements in presenting the results of the study.</p> <p style="text-align: center;">Хорошо</p> <p>The student knows the requirements for the international presentation of the research data in a written form.</p> <p>The student is able to use bibliographic managers to provide references in various formats.</p> <p>The student can use the results of smart research to justify and illustrate scientific papers but cannot meet the international requirements in presenting the results of the study.</p> <p style="text-align: center;">Отлично</p> <p>The student knows the requirements for the international presentation of the research data in a written form.</p> <p>The student is able to use bibliographic managers to provide references in various formats.</p> <p>The student can use the results of smart research to justify and illustrate scientific papers and present the results of the study in accordance with international requirements.</p>
--	---	---

ПК.3

Способен проводить самостоятельные научные исследования по актуальной проблематике в области лингвистики, переводоведения и межкультурной коммуникации

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ПК.3.2 Применяет современные методики поиска, анализа и обработки материала исследования на основе системного представления о динамике развития переводоведения,</p>	<p>Ability to apply modern methods of searching, analyzing and processing research material based on a systemic view of the dynamics of development of translation studies, linguistics, and intercultural communication</p>	<p style="text-align: center;">Неудовлетворительно</p> <p>The required skills have not been developed.</p> <p style="text-align: center;">Удовлетворительно</p> <p>The student knows modern tools for textual data processing;</p> <p>The student is not able to use programing tools and set workflows to solve one of the following tasks: to automate parsing, textual data parsing, preparation, and processing;</p> <p>The student cannot apply modern tools to estimate the results of computer-aided data</p>

<p>лингвистики и межкультурной коммуникации</p>		<p style="text-align: center;">Удовлетворительно</p> <p>processing.</p> <p style="text-align: center;">Хорошо</p> <p>The student knows modern tools for textual data processing;</p> <p>The student is not able to use programming tools and set workflows to solve one of the following tasks: to automate parsing, textual data parsing, preparation, and processing;</p> <p>The student applies modern tools to estimate the results of computer-aided data processing.</p> <p style="text-align: center;">Отлично</p> <p>The student knows modern tools for textual data processing;</p> <p>The student is able to use programming tools and set workflows to automate parsing, textual data parsing, preparation, and processing;</p> <p>The student applies modern tools to estimate the results of computer-aided data processing.</p>
---	--	--

Оценочные средства

Вид мероприятия промежуточной аттестации : Экзамен

Способ проведения мероприятия промежуточной аттестации : Защищаемое контрольное мероприятие

Продолжительность проведения мероприятия промежуточной аттестации :
время отводимое на доклад 4

Показатели оценивания

<p>The practical assignment including formulating research questions based on provided scenarios, designing a research study on a chosen linguistic topic, collecting and analyzing data using appropriate methods has not been done.</p>	<p>Неудовлетворительно</p>
<p>Research questions have been formulated based on provided scenarios. A research study on a chosen linguistic topic has been designed. Data have been collected and analyzed using appropriate methods. Neither a research paper based on the findings and supplied with supporting data and visuals has been written nor research results have not been communicated through a presentation containing supporting data and visualization means.</p>	<p>Удовлетворительно</p>
<p>Research questions have been formulated based on provided scenarios. A research study on a chosen linguistic topic has been designed. Data have been collected and analyzed using appropriate methods. Either a research paper based on the findings and supplied with supporting data and visuals has been written or research results have been communicated through a presentation containing supporting data and visualization means.</p>	<p>Хорошо</p>
<p>Research questions have been formulated based on provided scenarios.</p>	<p>Отлично</p>

A research study on a chosen linguistic topic has been designed.
Data have been collected and analyzed using appropriate methods.
A research paper based on the findings has been written and supplied with supporting data and visuals.
Research results have been communicated through a presentation containing supporting data and visualization means.

Отлично