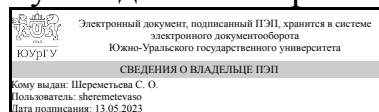


УТВЕРЖДАЮ:  
Руководитель направления



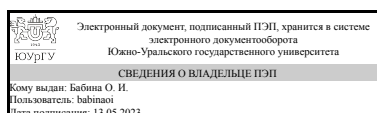
С. О. Шереметьева

## РАБОЧАЯ ПРОГРАММА

**дисциплины 1.Ф.01 Лингвистические базы знаний и языковые ресурсы для направления 45.04.03 Фундаментальная и прикладная лингвистика**  
**уровень** Магистратура  
**форма обучения** очная  
**кафедра-разработчик** Лингвистика и перевод

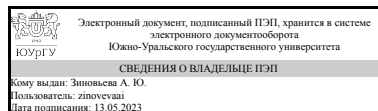
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 45.04.03 Фундаментальная и прикладная лингвистика, утверждённым приказом Минобрнауки от 12.08.2020 № 993

Зав.кафедрой разработчика,  
к.филол.н., доц.



О. И. Бабина

Разработчик программы,  
к.филол.н., преподаватель



А. Ю. Зиновьева

## 1. Цели и задачи дисциплины

Цель дисциплины — ознакомление с принципами построения, функционирования и использования существующих лингвистических баз данных, баз знаний и прочих языковых ресурсов. В ходе дисциплины решаются такие задачи, как овладение терминологическим аппаратом компьютерной лингвистики, связанным с построением, функционированием и использованием лингвистических баз данных, баз знаний и прочих языковых ресурсов; изучение характеристик лингвистических баз данных, баз знаний и прочих языковых ресурсов; ознакомление с основными принципами их разработки и использования.

## Краткое содержание дисциплины

Лингвистические ресурсы как основа лингвистических исследований. Виды и типы лингвистических баз данных. Корпусы текстов. Словари, лексиконы, терминологические банки данных. Онтологии и ресурсы онтологического типа.

## 2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-1 Способен осуществлять проектирование и разработку лингво-информационных ресурсов	Знает: виды лингво-информационных ресурсов и их характеристики; принципы, методы и средства проектирования и построения архитектуры лингво-информационных ресурсов; типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке лингво-информационных ресурсов Умеет: использовать существующие типовые решения и шаблоны лингво-информационных ресурсов; применять принципы, методы и средства проектирования лингво-информационных ресурсов Имеет практический опыт: разработки, изменения архитектуры лингво-информационных ресурсов; проектирования структур данных, баз данных и баз знаний
ПК-2 Способен управлять получением, хранением и обработкой языковых данных для решения профессиональных задач	Знает: основные методы, способы и средства управления получением, хранением и обработкой языковых данных; источники языковых данных; методы и средства извлечения и преобразования языковых данных; правовые основы сбора и использования языковых данных для решения профессиональных задач Умеет: взаимодействовать с пользователями информации и поставщиками языковых данных; пользоваться методами и средствами поиска, извлечения, хранения, передачи, преобразования и обработки языковых данных для решения профессиональных задач Имеет практический опыт: поиска, извлечения,

	хранения, передачи, преобразования и обработки языковых данных для решения профессиональных задач
--	---

### 3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Нет	Не предусмотрены

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

### 4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 2 з.е., 72 ч., 36,25 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		2	
Общая трудоёмкость дисциплины	72	72	
<i>Аудиторные занятия:</i>	32	32	
Лекции (Л)	0	0	
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	32	32	
Лабораторные работы (ЛР)	0	0	
<i>Самостоятельная работа (СРС)</i>	35,75	35,75	
Подготовка к зачету	5,75	5,75	
Оформление отчетов по практическим работам	30	30	
Консультации и промежуточная аттестация	4,25	4,25	
Вид контроля (зачет, диф.зачет, экзамен)	-	зачет	

### 5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Лингвистические ресурсы как основа лингвистических исследований	6	0	6	0
2	Виды и типы лингвистических баз данных	4	0	4	0
3	Корпусы текстов	4	0	4	0
4	Словари, лексиконы, терминологические банки данных	8	0	8	0
5	Онтологии и ресурсы онтологического типа	10	0	10	0

## 5.1. Лекции

Не предусмотрены

## 5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во часов
1	1	База данных: общие характеристики и принципы построения. Лингвистические базы данных и лингвистические ресурсы как смежные понятия. Типы лингвистических ресурсов. Международные стандарты представления лингвистических данных в электронном формате.	2
2	1	Первичные и вторичные лингвистические данные. Средства обработки речевого материала.	2
3	1	Автоматический анализ и синтез звучащей речи. Автоматическое распознавание текста. Автоматический анализ и синтез текста.	2
4	2	Классификационные базы данных. Тезаурусные базы данных, базы данных предметных рубрик, словарные базы данных. Отражение в электронных ресурсах информации о различных языках. Принципы работы и возможности поиска в следующих базах: A world-wide survey of affix borrowing, The World Loanword Database, The World Atlas of Language Structures.	2
5	2	Классификации фонетических баз данных по материалу и по назначению. Методы сбора и включения материала в фонетические базы данных. Основные принципы работы со звуковыми редакторами. Обзор современных фонетических баз данных.	2
6	3	Структура и виды корпусов текстов. Национальный корпус русского языка. Национальный корпус английского языка (британский вариант). Национальный корпус английского языка (американский вариант). Другие национальные корпуса. Принципы работы с корпусами текстов. Корпус-менеджеры.	2
7	3	Принципы создания параллельных и псевдопараллельных корпусов предметной области. Виды корпусной разметки.	2
8	4	Типы электронных словарей и общие принципы построения. Общий обзор существующих электрон-ных словарей. Перспективы развития электронной лексикографии.	2
9	4	Терминологические банки данных. Управление операциями с терминологией. Автоматическое извлечение терминов из текстов на разных языках.	2
10–11	4	Принципы разработки словарей предметной области.	4
12	5	Компьютерные тезаурусы. Тезаурусы для русского, английского языков. Надстройки для компьютерных тезаурусов.	2
13	5	Формальные онтологии: структура и классификация. Подходы к разработке формальных онтологий. Языки представления и редакторы формальных онтологий. Методы автоматизации построения формальных онтологий. Современные онтологии и ресурсы онтологического типа.	2
14	5	Принципы работы с существующими тезаурусами и онтологиями.	2
15–16	5	Разработка онтологии предметной области.	4

## 5.3. Лабораторные работы

Не предусмотрены

## 5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к зачету	1. Шереметьева, С. О. Linguistic models and tools for processing patent claims [Текст] monograph : на англ. яз. S. O. Sheremetyeva ; South Ural State University, Dep. of Linguistics and International Relations ; ЮУрГУ. - Chelyabinsk: SUSU publishing centre, 2017 2. Шипицина, Л. Ю. Информационные технологии в лингвистике : учебное пособие / Л. Ю. Шипицина. — 3-е изд., стер. — Москва : ФЛИНТА, 2017. — 126 с. — ISBN 978-5-9765-1431-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/119463">https://e.lanbook.com/book/119463</a> (дата обращения: 23.01.2022). — Режим доступа: для авториз. пользователей. 3. Лукашевич, Н. В. Тезаурусы в задачах информационного поиска : учебное пособие / Н. В. Лукашевич. — Москва : МГУ имени М.В.Ломоносова, 2011. — 512 с. — ISBN 978-5-211-05926-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/10143">https://e.lanbook.com/book/10143</a> (дата обращения: 23.01.2022). — Режим доступа: для авториз. пользователей.	2	5,75
Оформление отчетов по практическим работам	-	2	30

## 6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

### 6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	2	Текущий контроль	Тест	1	10	За каждый правильный ответ на вопрос студент получает 1 балл. Максимальное количество баллов — 10.	зачет
2	2	Текущий контроль	Презентация	2	8	Содержание доклада соответствует заявленной теме — 1 балл	зачет

						<p>Проведен сравнительно-сопоставительный анализ двух или трех баз данных — 1 балл</p> <p>Приведено полное описание каждой из баз данных — 1 балл</p> <p>Выделены общие черты баз данных — 1 балл</p> <p>Выделены различающиеся черты баз данных — 1 балл</p> <p>Приведены и обоснованы преимущества баз данных — 1 балл</p> <p>Приведены и обоснованы недостатки баз данных — 1 балл</p> <p>Описаны возможные способы применения баз данных — 1 балл</p>	
3	2	Текущий контроль	Практическая работа 1	3	6	<p>Собран корпус текстов на русском языке — 1 балл</p> <p>Собран корпус текстов на иностранном языке — 1 балл</p> <p>Корпусы параллельны или псевдопараллельны — 1 балл</p> <p>Объем корпусов составляет 5 тыс. словоупотреблений — 1 балл</p> <p>Оба корпуса соответствуют тематике — 1 балл</p> <p>Оба корпуса репрезентативны — 1 балл</p>	зачет
4	2	Текущий контроль	Практическая работа 2	3	8	<p>Составлен частотный словарь на русском языке — 1 балл</p> <p>Составлен частотный словарь на иностранном языке — 1 балл</p> <p>В словарях отсутствуют единицы, нерелевантные для предметной области — 1 балл</p> <p>Частота многозначных единиц разделена — 1 балл</p> <p>Создан двуязычный лексикон — 1 балл</p> <p>Выделены релевантные имена собственные — 1 балл</p> <p>Создан двуязычный ономастикон — 1 балл</p> <p>Синонимичные названия сгруппированы — 1 балл</p>	зачет
5	2	Текущий контроль	Практическая работа 3	3	7	<p>Разработанный ресурс является онтологией (не тезаурусом) предметной области — 1 балл</p> <p>Выделено не менее 15 базовых концептов — 1 балл</p> <p>Выделено не менее четырех отношений между концептами — 1 балл</p> <p>В онтологии отсутствуют концепты, не связанные друг с другом — 1 балл</p> <p>Выделено не менее двух атрибутов концептов — 1 балл</p> <p>Фреймы концептов заполнены корректно — 1 балл</p> <p>Создана по крайней мере одна аксиома</p>	зачет

						— 1 балл	
6	2	Промежуточная аттестация	Зачетное задание	-	7	Объем основной части отчета соответствует требованиям — 1 балл Выбор предметной области обоснован — 1 балл Ход работы описан — 1 балл Проблемы, возникшие при разработке онтологии, описаны — 1 балл Приведены способы решения проблем, возникших при разработке — 1 балл Выводы по проделанной работе сделаны — 1 балл Имеется приложение, содержащее полное описание всех компонентов онтологии — 1 балл	зачет
7	2	Бонус	Доклад на конференции по теме дисциплины	-	5	В случае участия с докладом по теме дисциплины на конференции студент получает 5 % бонусного рейтинга.	зачет
8	2	Бонус	Публикация статьи в рецензируемом журнале по теме дисциплины	-	10	В случае публикации статьи по теме дисциплины в рецензируемом журнале студент получает 10 % бонусного рейтинга.	зачет

## 6.2. Процедура проведения, критерии оценивания

Не предусмотрены

## 6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ KM							
		1	2	3	4	5	6	7	8
ПК-1	Знает: виды лингво-информационных ресурсов и их характеристики; принципы, методы и средства проектирования и построения архитектуры лингво-информационных ресурсов; типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке лингво-информационных ресурсов	+	+	+	+	+	+	+	+
ПК-1	Умеет: использовать существующие типовые решения и шаблоны лингво-информационных ресурсов; применять принципы, методы и средства проектирования лингво-информационных ресурсов	+	+	+	+	+	+	+	+
ПК-1	Имеет практический опыт: разработки, изменения архитектуры лингво-информационных ресурсов; проектирования структур данных, баз данных и баз знаний			+	+	+	+	+	+
ПК-2	Знает: основные методы, способы и средства управления получением, хранением и обработкой языковых данных; источники языковых данных; методы и средства извлечения и преобразования языковых данных; правовые основы сбора и использования языковых данных для решения профессиональных задач	+	+	+	+	+	+	+	+
ПК-2	Умеет: взаимодействовать с пользователями информации и поставщиками языковых данных; пользоваться методами и средствами поиска, извлечения, хранения, передачи, преобразования и обработки языковых данных для решения профессиональных задач			+	+	+	+	+	+
ПК-2	Имеет практический опыт: поиска, извлечения, хранения, передачи, преобразования и обработки языковых данных для решения			+	+	+	+	+	+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

## 7. Учебно-методическое и информационное обеспечение дисциплины

### Печатная учебно-методическая документация

#### а) основная литература:

1. Шереметьева, С. О. Linguistic models and tools for processing patent claims [Текст] monograph : на англ. яз. S. O. Sheremetyeva ; South Ural State University, Dep. of Linguistics and International Relations ; ЮУрГУ. - Chelyabinsk: SUSU publishing centre, 2017

#### б) дополнительная литература:

Не предусмотрена

#### в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

#### г) методические указания для студентов по освоению дисциплины:

1. Методические указания

#### из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания

### Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Дополнительная литература	Электронно-библиотечная система издательства Лань	Лукашевич, Н. В. Тезаурусы в задачах информационного поиска : учебное пособие / Н. В. Лукашевич. — Москва : МГУ имени М.В.Ломоносова, 2011. — 512 с. — ISBN 978-5-211-05926-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/10143">https://e.lanbook.com/book/10143</a> (дата обращения: 23.01.2023). — Режим доступа: для авториз. пользователей.
2	Основная литература	Электронно-библиотечная система издательства Лань	Шипицина, Л. Ю. Информационные технологии в лингвистике : учебное пособие / Л. Ю. Шипицина. — 3-е изд., стер. — Москва : ФЛИНТА, 2017. — 126 с. — ISBN 978-5-9765-1431-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/119463">https://e.lanbook.com/book/119463</a> (дата обращения: 23.01.2023). — Режим доступа: для авториз. пользователей.

Перечень используемого программного обеспечения:

1. Microsoft-Windows(бессрочно)
2. Microsoft-Office(бессрочно)



Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

## 8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Зачет, диф. зачет	478 (1)	Оборудование и технические средства обучения: 1. Персональный компьютер студента – 16 шт. 2. Персональный компьютер преподавателя – 1 шт. 3. Монитор – 17 шт. 4. Мышь – 17 шт. 5. Клавиатура – 17 шт. 6. Микротелефонная гарнитура – 17 шт. 7. Интерактивная панель планшет – 1 шт. 8. Мультимедиапроектор – 1 шт. 9. Настенно-потолочный экран с электроприводом – 1 шт. 10. Активная акустическая система (аудиоколонки) – 1 шт. 11. Коммутатор – 1 шт. 12. Сетевой фильтр – 1 шт. Имущество: 1. Стул «Стандарт» – 44 шт. 2. Шкаф открытый – 1 шт. 3. Антресоль с замком – 1 шт. 4. Аудиторная доска – 1 шт. 5. Жалюзи вертикальные блэкаут – 3 шт. 6. Кронштейн (крепление для аудиоколонок) – 2 шт. 7. Кронштейн (крепление для проектора) – 1 шт. 8. Рабочий стол студента – 16 шт. 9. Рабочий стол преподавателя – 1 шт. 10. Специализированный стол для практических занятий – 1 шт.
Практические занятия и семинары	478 (1)	Оборудование и технические средства обучения: 1. Персональный компьютер студента – 16 шт. 2. Персональный компьютер преподавателя – 1 шт. 3. Монитор – 17 шт. 4. Мышь – 17 шт. 5. Клавиатура – 17 шт. 6. Микротелефонная гарнитура – 17 шт. 7. Интерактивная панель планшет – 1 шт. 8. Мультимедиапроектор – 1 шт. 9. Настенно-потолочный экран с электроприводом – 1 шт. 10. Активная акустическая система (аудиоколонки) – 1 шт. 11. Коммутатор – 1 шт. 12. Сетевой фильтр – 1 шт. Имущество: 1. Стул «Стандарт» – 44 шт. 2. Шкаф открытый – 1 шт. 3. Антресоль с замком – 1 шт. 4. Аудиторная доска – 1 шт. 5. Жалюзи вертикальные блэкаут – 3 шт. 6. Кронштейн (крепление для аудиоколонок) – 2 шт. 7. Кронштейн (крепление для проектора) – 1 шт. 8. Рабочий стол студента – 16 шт. 9. Рабочий стол преподавателя – 1 шт. 10. Специализированный стол для практических занятий – 1 шт.