

Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Оренбургский государственный университет»

Кафедра романской филологии и методики преподавания французского языка

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б1.Д.В.Э.2.2 Компьютерная обработка текста»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

45.03.02 Лингвистика

(код и наименование направления подготовки)

Перевод и переводоведение (французский язык, второй иностранный язык)
(наименование направленности (профиля) образовательной программы)

Квалификация

Бакалавр

Форма обучения

Очная

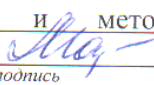
Год набора 2025

Рабочая программа дисциплины «Б1.Д.В.Э.2.2 Компьютерная обработка текста» рассмотрена и утверждена на заседании кафедры

Кафедра романской филологии и методики преподавания французского языка
наименование кафедры

протокол № 6 от «18» февраля 2025 г.

Заведующий кафедрой

Кафедра романской филологии и методики преподавания французского языка
наименование кафедры  И.Ю. Моисеева
подпись расшифровка подписи

Исполнители:

преподаватель

должность



подпись

Э.П. Аргунеев

расшифровка подписи

должность

подпись

расшифровка подписи

СОГЛАСОВАНО:

Председатель методической комиссии по направлению подготовки

45.03.02 Лингвистика

код наименование

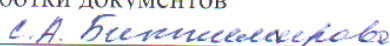

личная подпись

А.В. Павлова

расшифровка подписи

Заведующий отделом формирования фонда и научной обработки документов


личная подпись


расшифровка подписи

Уполномоченный по качеству факультета


личная подпись

расшифровка подписи

Т.В. Сапух

№ регистрации _____

© Аргунеев Э.П., 2025
© ОГУ, 2025

1 Цели и задачи освоения дисциплины

Цель (цели) освоения дисциплины: развитие способности применять автоматический лингвистический анализ и синтез текста, а также ознакомление студентов с лингвистическими аспектами общения с компьютером на естественном языке.

Задачи:

- владеть основами автоматического индексирования и реферирования текстов, предредактирования и постредактирования;
- владеть лингвистической и статистической терминологией в области современных методов обработки текста на естественном языке;
- уметь проводить статистическую обработку специальных текстов, машинный перевод и многоязыковую генерацию текстов;
- владеть основами системы извлечения текстовой информации, извлечения знаний из больших массивов информации;
- иметь навыки (приобрести опыт) работы с программным обеспечением для автоматического анализа текстов: морфологическими и синтаксическими анализаторами, конкордансами, классификации и тематического моделирования коллекций документов.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока Д «Дисциплины (модули)»

Пререквизиты дисциплины: *Б1.Д.Б.11 Информатика, Б1.Д.Б.12 Информационные технологии и программирование, Б1.Д.В.1 Теория перевода*

Постреквизиты дисциплины: *Отсутствуют*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
ПК*-2 Способен осуществлять подготовку к выполнению перевода	ПК*-2-В-2 Составляет необходимые глоссарии и терминологические базы, работает с готовыми глоссариями и терминологическими базами ПК*-2-В-6 Осуществляет техническую подготовку текста к переводу путем распознавания, набора, форматирования, извлечения из графики и конвертации в требуемые форматы с использованием доступных программных средств	Знать: методику технической и психологической подготовки к выполнению разных типов перевода; наиболее авторитетные словари (включая электронные), справочно-информационные базы, электронные ресурсы, поисковые системы; типы и принципы поисковых запросов. Уметь: составлять тезаурусы и глоссарии; пользоваться широким набором поисковых систем, словарей (включая электронные), справочно-

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
		информационных баз, электронных ресурсов; формулировать поисковые запросы и использовать поисковый и инструментальный и справочный аппарат; использовать подготовку к пере-воду в профессиональной деятельности. <u>Владеть:</u> стандартными методиками и навыками поиска информации в разных типах источников; разными видами чтения; навыками лингвистической и психологической подготовки к переводу; способностью оценивать уровень своей подготовки к переводу для ее корректировки и совершенствования.
ПК*-4 Способен редактировать и корректировать текст перевода, проверять его качество и соответствие переводческому заданию, оформлять текст перевода в соответствии с требованиями заказчика	ПК*-4-В-4 Адекватно применяет методы, процедуры и программные средства контроля качества перевода ПК*-4-В-5 Форматирует, конвертирует, верстает письменный текст перевода, восстанавливает исходное форматирование в соответствии с переводческим заданием; продуцирует устный текст перевода в соответствии с требованиями заказчика	<u>Знать:</u> .различные текстовые редакторы для оформления теста перевода; специфические возможности использования различных текстовых редакторов для оформления теста перевода; приемы редактирования и форматирования текста в различных текстовых редакторах. <u>Уметь:</u> использовать различные текстовые редакторы для создания, просмотра, форматирования и редактирования разных типов текстов. <u>Владеть:</u> навыками редактирования и форматирования разных типов текстов; навыками внедрения в текст сторонних объектов (шрифтов, таблиц, формул, графиков, фотографий, изображений, диаграмм).
ПК*-5 Способен использовать средства автоматизации переводческой деятельности и цифровые технологии для решения переводческих задач	ПК*-5-В-1 Корректно подготавливает исходный текст для работы в системах машинного перевода и САТ-программах в соответствии с требованиями ПК*-5-В-2 Выбирает соответствующие переводческой задаче цифровые технологии и средства автоматизации ПК*-5-В-3 Адекватно выполняет перевод в системах машинного	<u>Знать:</u> основные средства автоматизации переводческой деятельности; основные САТ – программы <u>Уметь:</u> использовать основные средства автоматического перевода; использовать основные САТ – программы <u>Владеть:</u> навыками работы с автоматизированными системами

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
	перевода и САТ-программах ПК*-5-В-4 Адекватно осуществляет постредактирование текста перевода ПК*-5-В-5 Выступает как участник и организатор переводческих проектов	перевода; навыками работы с САТ – программами.

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 3 зачетные единицы (108 академических часов).

Вид работы	Трудоемкость, академических часов	
	4 семестр	всего
Общая трудоёмкость	108	108
Контактная работа:	34,25	34,25
Практические занятия (ПЗ)	34	34
Промежуточная аттестация (зачет, экзамен)	0,25	0,25
Самостоятельная работа: - написание реферата (Р); - написание эссе (Э); - самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий); - подготовка к практическим занятиям; - подготовка к рубежному контролю и т.п.	73,75	73,75
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	зачет	

Разделы дисциплины, изучаемые в 4 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Системы автоматической обработки текста и лингвистический автомат	14		6		8
2	Подъязыки в системах машинного перевода	12		4		8
3	Теория конечных автоматов	12		4		8
4	Системы аннотирования и реферирования	14		4		10
5	Компьютерная (автоматическая) обработка текстов	14		4		10
6	Типы лингвистических словарей. Онлайн-словари и переводчики	14		4		10
7	Оформление текста перевода в компьютерном текстовом редакторе. Работа с текстовым редактором Microsoft Office	14		4		10
8	Параллельный корпус как инструмент	14		4		10

№ раздела	Наименование разделов	Количество часов			
		всего	аудиторная работа		
			Л	ПЗ	внеауд. работа
	переводчика				
	Итого:	108		34	74
	Всего:	108		34	74

4.2 Содержание разделов дисциплины

1. Системы автоматической обработки текста и лингвистический автомат. Лингвистический автомат. Человеко-машинная синергетика и самоорганизация в системах АОТ. Уровневое построение систем АОТ и ЛА. Опознавание языка документа и нормализация входного текста. Подблок опознавания формата текста и его частей, а также определение их жанровой и тематической принадлежности. Средства обеспечения живучести ЛА.

2. Подъязыки в системах машинного перевода. Подъязыки технической документации. Предредактирование. Постредактирование.

3. Теория конечных автоматов. Моделирование языка с помощью конечного автомата. Формальные языки. Недетерминированные конечные автоматы. Распознавание как поиск. Построение лексикона на основе конечного автомата.

4. Системы аннотирования и реферирования. Подходы и методы автоматического реферирования. Экстрактивные методы. Статистические методы. Логико-математические методы. Дистрибутивный метод. Метод содержательных аспектов. Метод текстовых связей. Методы с опорой на знания.

5. Компьютерная (автоматическая) обработка текстов. Классификация систем КОТ по выполняемым функциям. Системы первого типа (программы Pragma, Lingvo, Белазар, Promt, Socrat, Google Translate и др.). Системы генерации (синтеза) текстов по особому формальному описанию (программы SciGen, Textgen, VIRGO 4, Кобра 9.1). Трудности создания полноценно работающих с языком систем КОТ. Основные направления применения систем КОТ.

6. Типы лингвистических словарей. Онлайн словари и переводчики. Макроструктура словаря. Микроструктура словаря. Типология лингвистических словарей. Компьютерная обработка текста: описание моделей и направлений разработок. Статистика речи и автоматический анализ текста. Лингвистический мониторинг функционирования языка. Компьютерное моделирование языка и речи.

7. Оформление текста перевода в компьютерном текстовом редакторе. Работа с текстовым редактором Microsoft Office Microsoft Word. Ассистент слияния. Автотекст. Стили. Кадры. Шаблоны. Формы. Макрокоманды. Главный документ. Таблицы. Связь с другими документами.

8. Параллельный корпус как инструмент переводчика. Понятие корпусной лингвистики. Национальный корпус русского языка. Британский национальный корпус (British National Corpus, BNC). Лингвистические исследования на основе корпуса. Конкорданс как инструмент переводчика.

4.3 Практические занятия (семинары)

№ занятия	№ раздела	Тема	Кол-во часов
1-3	1	Системы автоматической обработки текста и лингвистический автомат	4
4-5	2	Подъязыки в системах машинного перевода	4
6-7	3	Теория конечных автоматов	4
8-9	4	Системы аннотирования и реферирования	4
10-11	5	Компьютерная (автоматическая) обработка текстов	4
12-13	6	Типы лингвистических словарей. Онлайн-словари и переводчики	4

№ занятия	№ раздела	Тема	Кол-во часов
14-15	7	Оформление текста перевода в компьютерном текстовом редакторе. Работа с текстовым редактором Microsoft Office	4
16-17	8	Параллельный корпус как инструмент переводчика	4
		Итого:	34

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1 Мишова, В.В. Мультимедийные технологии: практикум / В.В. Мишова; Министерство культуры Российской Федерации, Кемеровский государственный институт культуры, Институт информационных и библиотечных технологий. – Кемеров: Кемеровский государственный институт культуры (КемГИК), 2017. – 80 с.: ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=472682>

2 Нужнов, Е.В. Мультимедиа технологии: учебное пособие / Е.В. Нужнов; Южный федеральный университет, Инженерно-технологическая академия. – 2-е изд., перераб. и доп. – Ростов-на-Дону; Таганрог: Южный федеральный университет, 2017. – Ч. 1. Основы мультимедиа технологий. – 199 с.: ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=499905>

5.2 Дополнительная литература

1 Колокольникова, А.И. Информатика: учебное пособие : [16+] / А.И. Колокольникова. – 2-е изд., испр. и доп. – Москва; Берлин : Директ-Медиа, 2020. – 289 с.: ил., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=596690>

2. Овчинникова, И. Г. Компьютерное моделирование вербальной коммуникации: учеб.-метод. пособие / И. Г. Овчинникова, И. А. Угланова. - М.: Флинта; : Наука, 2009. - 136 с. - ISBN 978-5-9765-0729-6. - ISBN 978-5-02-034802-8.

3. Онтологии и тезаурусы: модели, инструменты, приложения / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. – Москва: Интернет-Университет Информационных Технологий (ИНТУИТ) : Бином. Лаборатория знаний, 2009. – 173 с.: ил., табл., схем. – (Основы информационных технологий). – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=233056>

4. Семенов, А. Л. Современные информационные технологии и перевод: учеб. пособие для вузов / А. Л. Семенов. - М.: Академия, 2008. - 224 с. - ISBN 978-5-7695-4459-0.

5. Формализация исследовательских процедур анализа семантики языковых единиц: коллективная монография / М.В. Каменский, Т.Н. Ломтева, Н.С. Кабылкина и др.; под общ. ред. М.В. Каменского; Северо-Кавказский федеральный университет. – Ставрополь: Северо-Кавказский Федеральный университет (СКФУ), 2016. – 170 с.: ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=466913>

5.3 Периодические издания

Вестник компьютерных и информационных технологий: журнал. - М.: Агентство «Роспечать», 2024

5.4 Интернет-ресурсы

1. <https://openedu.ru/> - «Открытое образование».
2. <https://universarium.org/> - «Универсариум».
3. <https://www.lektorium.tv/> - «Лекториум».

5.5 Программное обеспечение, профессиональные базы данных и информационные справочные системы

Программное обеспечение

1. Пакет офисных приложений LibreOffice
2. Программная система для организации видео-конференц-связи Webinar.ru

Профессиональные базы данных и информационные справочные системы современных информационных технологий:

1. Академик [Электронный ресурс]: словари и энциклопедии. – Электрон. дан. - Москва [2000-2023]. – Режим доступа: <https://academic.ru/>
2. ГАРАНТ Платформа F1 [Электронный ресурс]: справочно-правовая система. / Разработчик ООО НПП «ГАРАНТ-Сервис», 119992, Москва, Воробьевы горы, МГУ, [1990–2023]. – Режим доступа в сети ОГУ для установки системы: \\fileserver1\GarantClient\garant.exe
3. КонсультантПлюс [Электронный ресурс]: электронное периодическое издание справочная правовая система. / Разработчик ЗАО «Консультант Плюс», [1992–2023]. – Режим доступа к системе в сети ОГУ для установки системы: \\fileserver1\CONSULT\cons.exe
4. Национальный корпус русского языка [Электронный ресурс]: онлайн-корпус русских текстов – Электрон. дан. - Москва [2003-2023]. – Режим доступа: ruscorpora.ru

6 Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения занятий семинарского типа, для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Аудитории оснащены комплектами ученической мебели, техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Помещение для самостоятельной работы обучающихся оснащены компьютерной техникой, подключенной к сети «Интернет», и обеспечением доступа в электронную информационно-образовательную среду ОГУ.