



федеральное государственное автономное образовательное учреждение высшего образования
Первый Московский государственный медицинский университет имени И.М. Сеченова
Министерства здравоохранения Российской Федерации
(Сеченовский Университет)

Утверждено
Ученый совет ФГАОУ ВО Первый МГМУ
им. И.М. Сеченова Минздрава России
(Сеченовский Университет)
«14» июня 2022
протокол №5

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Информационный поиск

основная профессиональная Высшее образование - бакалавриат - программа бакалавриата

45.00.00 Языкознание и литературоведение

45.03.04 Интеллектуальные системы в гуманитарной сфере

Интеллектуальные системы в лингвистике и медицине

Цель освоения дисциплины Информационный поиск

Цель освоения дисциплины: участие в формировании следующих компетенций:

ОПК-2; Способен к профессиональному росту и самосовершенствованию в области гуманитарных, социальных и лингвистических наук, а также в сфере техники и технологии информатики;

ОПК-5; Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности

Требования к результатам освоения дисциплины.

Изучение дисциплины направлено на формирование у обучающихся следующих компетенций:

п/№	Код компетенции	Содержание компетенции и (или ее части)	Индикаторы достижения компетенций:			
			Знать	Уметь	Владеть	Оценочные средства
1	ОПК-2	Способен к профессиональному росту и самосовершенствованию в области гуманитарных, социальных и лингвистических наук, а также в сфере техники и технологии информатики;	Методы доступа к информационным ресурсам	Пользоваться современными справочными и библиотечными системами и системами дистанционного поиска информации	Практическим опытом работы с поисковыми машинами, справочными и библиотечными системами и системами	Информационный поиск



		еских наук, а также в сфере техники и технологии информатик и;		ного образования	дистанцион ного образования	
2	ОПК-5	Способен понимать принципы работы современны х информаци онных технологий и использоват ь их для решения задач профессион альной деятельност и	Существую щие в настоящее время программн ые комплексы реализации сложных алгоритмов	Анализиру вать программн ые средства; самостоятел ьно создавать прикладные программн ые средства на основе современны х информаци онных технологий и сетевых ресурсов	Методами и приемами создания прикладных программ в образовани и; методикой применения математиче ски сложных алгоритмов в современны х программн ых комплексах; приемами использова ния современны х программн ых комплексов; технологие й создания приложений	Информаци онный поиск

Разделы дисциплины и компетенции, которые формируются при их изучении

п/№	Код компетенции	Наименование раздела/темы дисциплины	Содержание раздела в дидактических единицах	Оценочные средства
1	ОПК-2	1. Основы информационного поиска 1.1 Информационный поиск	Понятия «информация», и образ документа» и	«документ», «поисковый и «поисковый
				Информацион ный поиск



4 000643 04502

		информационные потребности	образ запроса». Виды информационных потребностей и типы поисковых задач. Релевантность, точность и полнота поиска. Общая характеристика информационно-поисковых систем (библиотечные каталоги, электронные каталоги, полнотекстовые поисковые системы, научные базы данных, специальные лингвистические ресурсы). Этапы поискового процесса	
		1.2 Модели информационного поиска	Основные модели информационного поиска: булева модель, векторная модель, вероятностные модели. Индекс, терм, частота, взвешивание термов. Понятие tf-idf и основы ранжирования документов. Расширение запросов, фасетный поиск, фильтры и ограничители в поисковых системах	Информационный поиск
2	ОПК-5	2. Технологии и инструменты информационного поиска 2.1 Индексирование и языки запросов	Ручное и автоматическое индексирование. Ключевые слова, предметные рубрики, тезаурусы, онтологии. Метаданные и форматы описания документов (основные элементы). Формальные языки запросов: логические операторы (AND, OR, NOT), поиск по полям, фразовый поиск, усечение и маскирование, фильтры. Особенности синтаксиса запросов в разных системах.	Информационный поиск
		2.2	Поиск научной информации в	Информацион



		Информационный поиск гуманитарных лингвистических исследованиях	электронных в каталогах, библиографических и полнотекстовых базах (национальные и международные ресурсы). Поиск по лингвистическим корпусам и словарным базам. Стратегии поиска для исследовательских задач (подбор литературы, поиск примеров языковых явлений, поиск параллельных текстов). Оценка надежности источников. Основы использования библиографических менеджеров и сервисов для организации найденной информации.	ный поиск
3	ОПК-5	3. 3.Языковые и программные средства информационного поиск 3.1 Базовая обработка текстовых данных для информационного поиска (Python, регулярные в	Представление текстов в виде данных. Понятия токенизации, нормализации, стоп-слов (на интуитивном уровне, без углубленной морфологии). Введение в использование Python для решения задач информационного поиска гуманитарной направленности: структура простого Python-скрипта; запуск кода (ин-терактивная среда / ноутбук); чтение текстовых файлов и вы-вод результатов (подсчёт строк, подсчёт вхождений слов); модуль re в Python: понятие регулярного выражения, шаблон, совпадение; простые регулярные выражения для поиска слов, дат, числовых значений, терминов в текстах; поиск всех вхождений по шаблону, извлечение фрагментов текста по регулярному	Информационный поиск



		выражению; сопоставление «чистого» поискового запроса и гегах-запроса: преимущества, ограничения, риски.	
	3.2 Программные средства информационного поиска: Python и регулярные выражения для работы с	Использование Python для автоматизации поисковых задач: чтение коллекций текстовых файлов (корпус мини-масштаба); фильтрация файлов по имени и содержимому (поиск релевантных документов); применение регулярных выражений для поиска паттернов в корпусе (лексические, морфологические, структурные паттерны); построение простейшей «поисковой статистики»: подсчёт числа документов, содержащих выражение/паттерн; подсчёт частоты вхождений; извлечение контекстов (конкордансы) с помощью регулярных выражений; сравнительный анализ: чем отличается ручной поиск в интерфейсе (корпус, поисковая система) от автоматизированного поиска через Python-скрипт; обсуждение реплицируемости и прозрачности поиска в исследовательской работе (проектирование эксперимента, фиксация запросов и кода); знакомство с базовыми библиотеками Python, применимыми к информационному поиску и анализу текстов; знакомство с nltk или аналогов	Информационный поиск

Виды учебной работы

Вид учебной работы	Трудоемкость		Трудоемкость по семестрам (Ч)
	объем в зачетных единицах	Объем в часах (Ч)	Семестр 4



	(ЗЕТ)		
Контактная работа, в том числе		60	60
Консультации, аттестационные испытания (КАтт) (Экзамен)		4	4
Лекции (Л)		16	16
Лабораторные практикумы (ЛП)			
Практические занятия (ПЗ)		40	40
Клинико-практические занятия (КПЗ)			
Семинары (С)			
Работа на симуляторах (РС)			
Самостоятельная работа студента (СРС)		30	30
ИТОГО	3	90	90

Содержание дисциплины (модуля) по видам занятий

Лекционные занятия

№ раздел а	Наименование раздела дисциплины (модуля)	Тема лекции	Применение ЭО и ДОТ	Объем, час.
1	Основы информационного поиска	Информационный поиск и информационные потребности	Размещено в Информационной системе «Университет-Обучающийся»	4
1	Основы информационного поиска	Модели информационного поиска	Размещено в Информационной системе «Университет-Обучающийся»	4
2	Технологии и инструменты информационного поиска	Индексирование и языки запросов	Размещено в Информационной системе «Университет-Обучающийся»	2
2	Технологии и инструменты информационного поиска	Информационный поиск в гуманитарных и лингвистических исследованиях	Размещено в Информационной системе «Университет-Обучающийся»	2
3	3.Языковые и программные средства информационного поиска	Базовая обработка текстовых данных для информационного поиска (Python, регулярные выражения)	Размещено в Информационной системе «Университет-Обучающийся»	2
3	3.Языковые и программные средства информационного поиска	Программные средства информационного поиска: Python и регулярные выражения для работы с	Размещено в Информационной системе «Университет-Обучающийся»	2



Практические занятия

№ раздела	Наименование раздела дисциплины (модуля)	Тема	Применение ЭО и ДОТ	Объем, час.
1	Основы информационного поиска	Информационный поиск и информационные потребности	Размещено в Информационной системе «Университет-Обучающийся»	7
1	Основы информационного поиска	Модели информационного поиска	Размещено в Информационной системе «Университет-Обучающийся»	7
2	Технологии и инструменты информационного поиска	Индексирование и языки запросов	Размещено в Информационной системе «Университет-Обучающийся»	7
2	Технологии и инструменты информационного поиска	Информационный поиск в гуманитарных и лингвистических исследованиях	Размещено в Информационной системе «Университет-Обучающийся»	7
3	3.Языковые и программные средства информационного поиска	Базовая обработка текстовых данных для информационного поиска (Python, регулярные выражения)	Размещено в Информационной системе «Университет-Обучающийся»	6
3	3.Языковые и программные средства информационного поиска	Программные средства информационного поиска: Python и регулярные выражения для работы с	Размещено в Информационной системе «Университет-Обучающийся»	6

Самостоятельная работа студента

№ раздела	Наименование раздела дисциплины (модуля)	Тема занятия	Вид СРС	Объем, час.
1	Основы информационного поиска	Информационный поиск и информационные потребности		5
1	Основы информационного поиска	Модели информационного поиска	Формулировка информационных потребностей и построение поисковых запросов в различных системах	5
2	Технологии и инструменты информационного поиска	Индексирование и языки запросов	Практика работы с моделями информационного поиска (булев поиск, tf-idf, оценка релевантности)	5
2	Технологии и инструменты информационного поиска	Информационный поиск в	Практика поиска в научных	5



	инструменты информационного поиска	гуманитарных и лингвистических исследованиях	базах данных, электронных библиотеках, корпусах и словарях; оформление результатов поиска для исследовательской работы	
3	3.Языковые и программные средства информационного поиска	Базовая обработка текстовых данных для информационного поиска (Python, регулярные в	Практикум по Python и регулярным выражениям: поиск паттернов в текстах и мини-корпусах, извлечение контекстов	5
3	3.Языковые и программные средства информационного поиска	Программные средства информационного поиска: Python и регулярные выражения для работы с	Практикум по Python и регулярным выражениям: поиск паттернов в текстах и мини-корпусах, извлечение контекстов	5

Учебно-методическое и информационное обеспечение дисциплины

Перечень основной литературы

№	Наименование согласно библиографическим требованиям
1	Нугуманова, А. Б. Автоматизированная обработка текстовых массивов : учебник и практикум для вузов / А. Б. Нугуманова. — 2-е изд. — Москва : Издательство Юрайт, 2024. — 82 с. — (Высшее образование). — ISBN 978-5-534-20738-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/558668
2	Пшехотская, Е. А. Информационный поиск : учебное пособие / Е. А. Пшехотская, Ю. Н. Филиппович. — Москва : Московский Политех, 2023. — 108 с. — ISBN 978-5-2760-2830-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/482786

Перечень дополнительной литературы

№	Наименование согласно библиографическим требованиям
1	Войтович Н. В. Информационный поиск в гуманитарных науках : учебное пособие
2	Мандельброт В. Теория информации и психолингвистика: теория частот слов // Математические методы в социальных науках. М., 1993. [С. 316–338.]
3	Алексеев П. М. Статистическая лексикография. Л., 1980
4	Баранов А.Н. Введение в прикладную лингвистику. М., 2001
5	www.ruscorpora.ru Национальный корпус русского языка
6	www.bnc.com British National Corpus (Британский национальный корпус)

Перечень электронных образовательных ресурсов

№	Наименование ЭОР	Ссылка
---	------------------	--------



1	Подготовка к итоговой аттестации_ИНФОРМАЦИОННЫЙ ПОИСК_ИСвГС (Б)	Размещено в Информационной системе «Университет-Обучающийся»
2	Информационный поиск	Размещено в Информационной системе «Университет-Обучающийся»
3	Информационный поиск ТМ	Размещено в Информационной системе «Университет-Обучающийся»

Материально-техническое обеспечение дисциплины

№ п/п	№ учебных аудиторий и объектов для проведения занятий	Адрес учебных аудиторий и объектов для проведения занятий	Наименование оборудованных учебных кабинетов, объектов для проведения практических занятий, объектов физической культуры и спорта с перечнем основного оборудования
1	403	123242, г. Москва, ул. Садовая-Кудринская, д. 3, стр. 1	
2	405	123242, г. Москва, ул. Садовая-Кудринская, д. 3, стр. 1	
3	305	123242, г. Москва, ул. Садовая-Кудринская, д. 3, стр. 1	

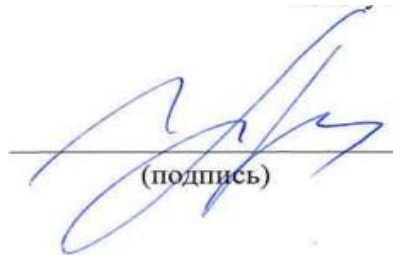


Рабочая программа дисциплины разработана Институтом иностранных языков для профессиональных целей

Принята на заседании Институтом иностранных языков для профессиональных целей

от «22» января 2025 г., протокол № 5

Директор
Института иностранных языков
для профессиональных целей



(подпись)

И.Ю.Марковина
(фамилия, инициалы)

Одобрена Центральным методическим советом
от «31» января 2025 г., протокол № 2