МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Санкт-Петербургский государственный экономический университет»

УТВЕРЖДАЮ
Проректор по учебной и
методической работе
_____/ Шубаева В.Г./
«17» 06 2019 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

МЕТОДЫ И СРЕДСТВА ИНФОРМАЦИОННОГО ПОИСКА

Направление подготовки	01.03.02 Прикладная математика и информатика
Направленность (профиль) программы	Прикладная математика и информатика в экономике и управлении
Уровень высшего образования	бакалавриат
Форма обучения	очная

Составители:
_____/ к.э.н., доцент Гниденко И.Г.
_____/ старший преподаватель Федоров Д.Ю.
_____/ к.т.н., руководитель аналитического управления ООО «Дрогери ритейл» Карахтанова А.Б.

Санкт-Петербург
2019

СОДЕРЖАНИЕ

1. ЦЕЛЬ ДИСЦИПЛИНЫ 3
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ 3
3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ 3
4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ 4
5. СОДЕРЖАНИЕ РАЗДЕЛОВ И ТЕМ ДИСЦИПЛИНЫ 4
6. ЗАНЯТИЯ СЕМИНАРСКОГО ТИПА 6
7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ 6
7.1. Методические указания для обучающегося по освоению дисциплины 6
7.2. Организация самостоятельной работы 7
8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ 7
9. РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ 8
9.1. Учебно-методическое и информационное обеспечение дисциплины 8
9.2. Материально-техническое обеспечение учебного процесса 9
10. ОСОБЕННОСТИ ОСВОЕНИЯ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ 10
11. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ 11

1. ЦЕЛЬ ДИСЦИПЛИНЫ

Цель дисциплины: представить обучающимся необходимые знания в области эксплуатации и разработки систем информационного поиска, умения и навыки использования современных средств разработки приложений.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина Б1.В.09 «Методы и средства информационного поиска» относится к части, формируемой участниками образовательных отношений Блока 1 учебного плана образовательной программы.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Планируемые результаты обучения по дисциплине, соотнесенные с установленными в образовательной программе индикаторами достижения компетенций, представлены в таблице 3.1.

Таблица 3.1 – Планируемые результаты обучения по дисциплине, соотнесенные с установленными в образовательной программе индикаторами достижения компетенций

Код и наименование компетенции выпускника	**Код и наименование индикаторов достижения компетенций**	**Планируемые результаты обучения по дисциплине**
1	**2**	**3**
ПК-1. Способен собрать информацию, организовать ее хранение и обработку на основе современных информационно-коммуникационных технологий, подготовить аналитический отчет.	ПК-1.1. Использует современные информационно-коммуникационные технологии для информационного поиска и сбора информации	***Знать***: современные методы и технологии организации информационно-аналитической деятельности, алгоритмы, лежащие в основе систем информационного поиска, перспективы развития систем информационного поиска ***Уметь:*** собирать, обрабатывать и анализировать данные, полученные с помощью систем информационного поиска ***Владеть***: навыками формирования аналитических отчетов с использованием современных программных решений
	ПК-1.3. Готовит аналитические отчеты, используя современные средства визуализации и стандарты оформления	*Знать*: современные средства формирования графических отчетов и сводных таблиц. *Уметь*: настраивать средства для сбора, анализа и визуализации данных. *Владеть*: навыками графического представления данных.

4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ

Трудоемкость дисциплины составляет 3 зачетные единицы, 108 часов.

Форма промежуточной аттестации: зачет в 7 семестре.

Распределение фонда времени по темам дисциплины по очной форме обучения представлено в таблице 4.1.

Таблица 4.1 – Распределение фонда времени по темам дисциплины

Номер и наименование тем	Объем дисциплины (ак. часы)			
	Контактная работа			СРО
	ЗЛТ	ПЗ	ЛР	
1	*2*	*3*	*4*	*5*
Тема 1. Форматы представления информации	2	-	-	10
Тема 2. Семантические сети. Проект Semantic Web	4	11	-	
Тема 3. Поисковые машины Интернета и технологии их создания. Понятие документа. Индексация документов	2	12	-	
Тема 4. Основные стандарты метаданных. Поиск документов различных форматов	2		-	10
Тема 5. Информационные потребности научного сообщества. Оценка эффективности поиска. Алгоритмы поиска «по аналогии»	4	11	-	
Тема 6. История возникновения тезаурусов и онтологий. Типы классификаторов. УДК. ГРНТИ	2		-	10
Тема 7. Методы и алгоритмы индексации документов	4	12	-	
Тема 8. Измерение близости между объектами. Основные алгоритмы кластеризации	2		-	10
Всего по дисциплине:	**22**	**46**		**40**

*ЗЛТ – занятия лекционного типа, ПЗ – все виды занятий семинарского типа, кроме лабораторных работ, ЛР – лабораторные работы, СРО – самостоятельная работа обучающегося

5. СОДЕРЖАНИЕ РАЗДЕЛОВ И ТЕМ ДИСЦИПЛИНЫ

Тема 1. Форматы представления информации

Основы теории семантической информации, определение информатики как науки о семантической информации, виды семантической информации, соотношение между понятиями «данные», «информация» и «знания».

Тема 2. Семантические сети. Проект Semantic Web

Онтологическое моделирование предметной области. Графовые базы данных и хранение данных в виде RDF-триплетов. Обращение к данным (БД) на логическом уровне. Графические интерфейсы в виде графов знаний. Плюсы и минусы использования семантических сетей.

Тема 3. Поисковые машины Интернета и технологии их создания. Понятие документа. Индексация документов

Технологии поиска. Поисковые инструменты. Поисковые машины (search engines). Каталоги (directories). Подборки ссылок. Базы данных адресов (addresses database). Поиск в архивах Gopher (Gopher archives). Система поиска FTP файлов (FTP Search). Система поиска в конференциях Usenet News. Системы мета-поиска. Системы поиска людей.

Тема 4. Основные стандарты метаданных. Поиск документов различных форматов

Различия между данными и метаданными. Классификация метаданных, классификация форматов метаданных, основные стандарты метаданных (GSDGM, ISO и т.д.), специальные программы для работы с метаданными, поиск документов различных форматов.

Тема 5. Информационные потребности научного сообщества. Оценка эффективности поиска. Алгоритмы поиска «по аналогии»

Информационные потребности научного сообщества, оценка эффективности поиска: полнота, точность, F-мера, базовые алгоритмы ранжирования, алгоритмы поиска «по аналогии»: алгоритм расширения выборки, метод n-грамм, хеширование по сигнатуре.

Тема 6. История возникновения тезаурусов и онтологий. Типы классификаторов. УДК. ГРНТИ

История возникновения тезаурусов и онтологий; типы классификаторов; типы онтологий; язык OWL; классификатор УДК; государственные рубрикатор ГРНТИ.

Тема 7. Методы и алгоритмы индексации документов

Взаимосвязь нечеткого поиска и индексации; частотный анализ; предварительная индексация документов; метод опорных векторов (SVM); алгоритм PrTFIDF; алгоритм «наивный Байес».

Тема 8. Измерение близости между объектами. Основные алгоритмы кластеризации

Измерение близости между объектами; цели кластеризации; основные методы кластеризации; евклидово расстояние; квадрат евклидова расстояния; расстояние городских кварталов; расстояние Чебышева; степенное расстояние.

6. ЗАНЯТИЯ СЕМИНАРСКОГО ТИПА

Таблица 6.1 – Практические занятия / Семинарские занятия / Лабораторные работы

№ темы	Тема занятия	Вид занятия / Оценочное средство
1	*2*	*3*
2	Поисковые информационные системы. Исследование эффективности поиска	ПЗ: Решение практической задачи 1
3	Разработка системы полнотекстового поиска с использованием Elasticsearch	ПЗ: Решение практической задачи 2
5	Настройка php клиента Elastica для системы полнотекстового поиска	ПЗ: Решение практической задачи 3
7	Настройка веб краулера Apache Nutch для системы полнотекстового поиска	ПЗ: Решение практической задачи 4

* ПЗ – практические занятия, СЗ – семинарские занятия, ЛР – лабораторные работы

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

7.1. Методические указания для обучающегося по освоению дисциплины

Для формирования четкого представления об объеме и характере знаний и умений, которыми надо будет овладеть по дисциплине в самом начале учебного курса обучающийся должен ознакомиться с учебно-методической документацией:

– рабочей программой дисциплины: с целями и задачами дисциплины, ее связями с другими дисциплинами образовательной программы, перечнем знаний и умений, которыми в процессе освоения дисциплины должен владеть обучающийся,

– порядком проведения текущего контроля успеваемости и промежуточной аттестации;

– графиком консультаций преподавателей кафедры.

Систематическое выполнение учебной работы на занятиях лекционных и семинарских типов, а также выполнение самостоятельной работы позволит успешно освоить дисциплину.

В процессе освоения дисциплины обучающимся следует:

– слушать, конспектировать излагаемый преподавателем материал;

– ставить, обсуждать актуальные проблемы курса, быть активным на занятиях;

– задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений;

– выполнять задания практических занятий полностью и установленные сроки.

При затруднениях в восприятии материала следует обратиться к основным литературным источникам. Если разобраться в материале не удалось, то обратится к лектору (по графику его консультаций) или к преподавателю на занятиях семинарского типа.

Обучающимся, пропустившим занятия (независимо от причин), не имеющим письменного решения задач или не подготовившимся к данному занятию, рекомендуется не позже, чем в 2 - недельный срок явиться на консультацию к преподавателю и отчитаться по теме.

7.2. Организация самостоятельной работы

Под самостоятельной работой обучающихся понимается планируемая работа обучающихся, направленная на формирование указанных компетенций, выполняемая во внеаудиторное время по заданию и при методическом руководстве преподавателя, без его непосредственного участия.

Методическое обеспечение самостоятельной работы при наличии обучающихся лиц с ограниченными возможностями представляется в формах, адаптированных к ограничениям их здоровья.

Виды самостоятельной работы по дисциплине представлены в таблице 7.2.1.

Таблица 7.2.1 – Организация самостоятельной работы обучающегося

№ темы	Вид самостоятельной работы
1	*2*
1	Самостоятельное задание: подготовить обзор определений термина «информация» Подготовка к ПЗ
4	Самостоятельное задание: подготовить обзор методов оценки эффективности поиска Подготовка к ПЗ
6	Самостоятельное задание: подготовить обзор типов классификаторов Подготовка к ПЗ
8	Самостоятельное задание: подготовить обзор метода опорных векторов Подготовка к ПЗ

Каждый вид СРО, указанный в таблице 7.2.1 обеспечен методическими материалами.

8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В рамках реализации дисциплины «*Методы и средства информационного поиска*» используются разнообразные образовательные технологии как традиционные, так и с применением активных и интерактивных методов обучения.

Все практические занятия проводятся в компьютерной аудитории в форме выполнения практических работ на компьютере. Активно используются обсуждения вариантов постановки задач, методов решения, возможных подходов к их компьютерной реализации. Совместно с обучающимися разрабатывается пошаговый план решения поставленной задачи. Решение *ИЗК* (индивидуальное задание, выполняемое на компьютере) реализуется средствами Python и PHP.

Проводится обсуждение адекватности вариантов полученного решения. Проводится совместное обсуждение корректности и возможности применения метода к другим практическим задачам.

9. РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

9.1. Учебно-методическое и информационное обеспечение дисциплины

Таблица 9.1.1 – Учебно-методическое обеспечение дисциплины

Библиографическое описание издания (автор, заглавие, вид, место и год издания, кол. стр.)	Основная/ дополнительная литература	Книгообеспеченность	
		Кол-во. экз. в библ. СПбГЭУ	Электронные ресурсы
1. Бессмертный И.А. Интеллектуальные системы : учебник и практикум / И.А. Бессмертный, А.Б. Нугуманова, А.В. Платонов. — Электрон. дан. — М. : Издательство Юрайт, 2019. — 243 с.	основная	–	ЭБС Юрайт
2. Интернет-маркетинг : учебник / О. Н. Жильцова [и др.] ; под общей редакцией О. Н. Жильцовой. — 2-е изд., перераб. и доп. — Электрон. дан. — Москва : Издательство Юрайт, 2019. — 301 с.	основная	–	ЭБС Юрайт
3. Станкевич Л.А. Интеллектуальные системы и технологии : учебник и практикум . — Электрон. дан. — М. : Издательство Юрайт, 2019. — 397 с.	основная	–	ЭБС Юрайт
4. Сысолетин Е.Г. Разработка интернет-приложений : учебное пособие / Е Г. Сысолетин, С.Д. Ростунцев ; под науч. ред. Л.Г. Доросинского. — Электрон. дан. — М. : Издательство Юрайт, 2019. — 90 с.	дополнительная	–	ЭБС Юрайт
5. Лаврищева Е.М. Программная инженерия и технологии программирования сложных систем : учебник — 2-е изд., испр. и доп. — Электрон. дан. — М. : Издательство Юрайт, 2019. — 432 с.	дополнительная	–	ЭБС Юрайт

Таблица 9.1.2 – Перечень современных профессиональных баз данных (СПБД)

№	Наименование СПБД
1	Электронная библиотека Grebennikon.ru – www.grebennikon.ru

2	Научная электронная библиотека eLIBRARRY – www.elibrary.ru
3	Научная электронная библиотека КиберЛеника – www.cyberleninka.ru
4	База данных ПОЛПРЕД Справочники – www.polpred.com
5	База данных OECD Books, Papers & Statistics на платформе OECD iLibrary – www.oecd-ilibrary.org

Таблица 9.1.3 – Перечень информационных справочных систем (ИСС)

№	Наименование ИСС
1	Справочная правовая система КонсультантПлюс (инсталлированный ресурс СПбГЭУ или www.consultant.ru)
2	Справочная правовая система «ГАРАНТ» (инсталлированный ресурс СПбГЭУ или www.garant.ru)
3	Информационно-справочная система «Кодекс» (инсталлированный ресурс СПбГЭУ или www.kodeks.ru)
4	Электронная библиотечная система BOOK.ru - www.book.ru
5	Электронная библиотечная система ЭБС ЮРАЙТ – www.urait.ru
6	Электронно-библиотечная система ЗНАНИУМ (ZNANIUM) – www.znanium.com
7	Электронная библиотека СПбГЭУ– opac.unecon.ru

9.2. Материально-техническое обеспечение учебного процесса

Для реализации данной дисциплины имеются специальные помещения для проведения занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Таблица 9.2.1 – Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства (ПО)

№	Наименование ПО
1	Microsoft Windows Professional (КОНТРАКТ № 244/20 «26» июня 2020 г.)
2	Microsoft Office Professional (КОНТРАКТ № 244/20 «26» июня 2020 г.)
3	7-Zip (freeware)
4	Wolfram Mathematica 12 (Сублицензионный договор № 328/20 от 20.08.2020г.)
5	Python 3.7.4 (freeware)

Таблица 9.2.2 – Перечень учебных аудиторий для проведения учебных занятий, оснащенных оборудованием и техническими средствами обучения

Наименование учебных аудиторий, перечень оборудования и технических средств обучения	**Адрес (местоположение) учебных аудиторий**
Ауд. 0007 Компьютерный класс (для проведения практических занятий, с применением вычислительной техники). Оборудован мультимедийным комплексом. Специализированная мебель и оборудование: Учебная мебель на 33 посадочных места, рабочее место	191023, г. Санкт-Петербург, ул. Канал Грибоедова, 30/32, литер «А», «Б», «Р»

преподавателя, доска меловая - 1 шт., доска маркерная на колесиках - 1 шт., вешалка стойка - 3 шт., жалюзи - 3 шт., Компьютер Intel Core i3 6100/ MSI H110M PRO-D/ ОЗУ DDR4 8GB 2400MHz/SSD SATA III 240Gb/Aerocool Qs-180 400W/Клавиатура + мышь Microsoft400 for Business/монитор Asus VS228DE - 24 шт., Мультимедийный проектор Тип 1 Optoma x 400 - 1 шт., Ноутбук HP 250 G6 1WY58EA - 2 шт. Наборы демонстрационного оборудования и учебно-наглядных пособий: мультимедийные приложения к лекционным курсам и практическим занятиям, интерактивные учебно-наглядные пособия.	
Ауд. 2032 Компьютерный класс (для проведения практических занятий, с применением вычислительной техники). Специализированная мебель и оборудование: Учебная мебель на 25 посадочных мест, рабочее место преподавателя (стол 1шт., кресло 1шт.), доска маркерная на колесиках 1 шт., маркерная доска на ножках 1шт., вешалки стойки 1шт., стол 2шт., стульев 4шт., доска обьявлений 1шт., жалюзи 2шт., Компьютер Intel I5-7400/16Gb/1Tb/ видеокарта NVIDIA GeForce GT 710/Монитор. DELL S2218H - 25 шт., Интерактивная доска SMARTB 680 - 1 шт., Шкаф телекоммуникационный настенный ЦМО ШРН-Э-6.650 - 1 шт., Коммутатор ProCurve Switch 2626 - 1 шт., Терминальная станция тонкий клиент в составе Sun Ray 2 client - 1 шт., Стойка для интерактивной доски 660x680 - 1 шт. Наборы демонстрационного оборудования и учебно-наглядных пособий: мультимедийные приложения к лекционным курсам и практическим занятиям, интерактивные учебно-наглядные пособия.	191023, г. Санкт-Петербург, ул. Канал Грибоедова, 30/32, литер «А», «Б», «Р»

10. ОСОБЕННОСТИ ОСВОЕНИЯ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

В целях освоения учебной программы дисциплины инвалидами и лицами с ограниченными возможностями здоровья Университет обеспечивает:

– для инвалидов и лиц с ограниченными возможностями здоровья по зрению: размещение в доступных для обучающихся, являющихся слепыми или слабовидящими, местах и в адаптированной форме справочной информации о расписании учебных занятий; присутствие ассистента, оказывающего обучающемуся необходимую помощь; выпуск альтернативных форматов методических материалов (крупный шрифт или аудиофайлы);

– для инвалидов и лиц с ограниченными возможностями здоровья по слуху: надлежащими звуковыми средствами воспроизведение информации;

– для инвалидов и лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата: возможность беспрепятственного доступа обучающихся в учебные помещения, туалетные комнаты и другие помещения кафедры, а также пребывание в указанных помещениях.

Обучающиеся из числа инвалидов и лиц с ОВЗ обеспечены печатными и (или) электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.

Образование обучающихся с ограниченными возможностями здоровья может быть организовано как совместно с другими обучающимися, так и в отдельных группах или в отдельных организациях.

11. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

Фонд оценочных средств для проведения аттестации уровня сформированности компетенций обучающихся по дисциплине оформляется отдельным документом и является приложением к рабочей программе дисциплины.