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

УТВЕРЖДАЮ
Проректор по учебной и методической работе
_____/Шубаева В.Г./
«20» июня 2018 г.

МЕТОДЫ И СРЕДСТВА ИНФОРМАЦИОННОГО ПОИСКА

Рабочая программа дисциплины

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

Составитель:
_____ к.т.н., рук. аналит. упр. ООО «Дрогери ритейл» Карахтанова А.Б.
_____ ст. преподаватель Федоров Д.Ю.

Санкт-Петербург
2018

СОДЕРЖАНИЕ

1. ЦЕЛЬ И ЗАДАЧИ 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 5
7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ 6
7.1. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩЕГОСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ 6
7.2. ОРГАНИЗАЦИЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ 6
8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ 7
9. РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ 7
9.1 УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ 7
9.2 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА 8
10. ОСОБЕННОСТИ ОСВОЕНИЯ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ 8
11. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ 9

1. ЦЕЛЬ И ЗАДАЧИ ДИСЦИПЛИНЫ

Цель (цели) освоения дисциплины: дать студентам необходимые знания в области эксплуатации и разработки систем информационного поиска, умения и навыки использования современных средств разработки приложений.

Задачи:

– знакомство с тенденциями развития систем информационного поиска;

– формирование умения использовать современные средства разработки систем информационного поиска;

– выработка практических навыков работы с инструментарием в области разработки приложений.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина Б1.В.ДВ..09.01 «Методы и средства информационного поиска» относится к выборным дисциплинам Блока 1 и является обязательной для освоения обучающимся после их выбора.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы представлены в таблице 3.1.

Таблица 3.1 – Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы

Код и наименование компетенции выпускника	Этапы формирования компетенций	Планируемые результаты обучения/индикаторы достижения компетенций (показатели освоения компетенции)
ОПК 2 способностью приобретать новые научные и профессиональные знания, используя современные образовательные и информационные технологии	Второй уровень (углубленный) (ОПК-2)-1	Знать: методы, инструменты и возможности поиска информации, анализа и работы с ней З2(III) (ОПК-2) Уметь: применять системы информационного поиска для приобретения новых знаний в области профессиональной деятельности, применять современные методы и технологии в информационно-аналитической деятельности У2(III) (ОПК-2) Владеть: навыками работы с современными средствами информационного поиска В2(III) (ОПК-2)
ПК 1 способностью собирать, обрабатывать и интерпретировать данные современных научных исследований, необходимые для формирования выводов по соответствующим научным исследованиям	ПК 1	Знать: современные методы и технологии организации информационно-аналитической деятельности, алгоритмы, лежащие в основе систем информационного поиска, перспективы развития систем информационного поиска З(VI) (ПК-1) Уметь: собирать, обрабатывать и анализировать данные, полученные с помощью систем информационного поиска У(VI) (ПК-1) Владеть: современными средствами информационного поиска В(VI) (ПК-1)

4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ

Трудоемкость дисциплины составляет 4 зачетные единицы, 144 часа.

Форма промежуточной аттестации: дифференцированный зачет (зачет с оценкой) - 8 семестр.

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

Таблица 4.1 – Распределение фонда времени по темам дисциплины (очная форма обучения)

Номер и наименование тем	Объем дисциплины (ак. часы)			
	Контактная работа			СРО
	ЗЛТ	ПЗ	ЛР	
1	2	3	4	5
Тема 1. Форматы представления информации.	5			20
Тема 2. Семантические сети. Проект Semantic Web.	5	10		
Тема 3. Поисковые машины Интернета и технологии их создания. Понятие документа. Индексация документов.	5			20
Тема 4. Основные стандарты метаданных. Поиск документов различных форматов.	5	10		
Тема 5. Информационные потребности научного сообщества. Оценка эффективности поиска. Алгоритмы поиска «по аналогии».	2			10
Тема 6. История возникновения тезаурусов и онтологий. Типы классификаторов. УДК. ГРНТИ.		10		10
Тема 7. Методы и алгоритмы индексации документов.		12		10
Тема 8. Измерение близости между объектами. Основные алгоритмы кластеризации.				10
Всего по дисциплине:	**22**	**42**		**80**

*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 прикладных задач
4	Разработка системы полнотекстового поиска с использованием Elasticsearch	ПЗ: Решение прикладных задач
6	Настройка php клиента Elastica для системы полнотекстового поиска	ПЗ: Решение прикладных задач КТ№1 Индивидуальное задание
7	Настройка веб краулера Apache Nutch для системы полнотекстового поиска	ПЗ: Решение прикладных задач КТ№2 Теоретический опрос

*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	Обзор определений термина «информация»
3	Обзор систем поиска людей
5	Обзор методов оценки эффективности поиска
6	Обзор типов классификаторов
7	Обзор метода опорных векторов

№ темы	Вид самостоятельной работы
8	Подготовка к теоретическому опросу

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

8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

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

Активные и интерактивные методы обучения:

К интерактивным методам обучения относится выполнение индивидуальных заданий на ПК в ходе практических занятий по темам 2, 4, 6, 7.

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 программного обеспечения (ПО)

№ п/п	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)

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.

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 (модуля).