



**Федеральное государственное автономное образовательное учреждение высшего
образования
«Национальный исследовательский университет
«Высшая школа экономики»**

Факультет Компьютерных наук
Департамент больших данных и информационного поиска

**Рабочая программа дисциплины
«Культура работы с данными»**

для образовательных программ
«Языки и литература Индии»,
«Язык и литература Ирана»,
«Библеистика и история древнего Израиля»
по направлению 58.03.01 «Востоковедение и африканистика»

уровень бакалавр

Разработчики программы

О.В. Дереза, преподаватель, odereza@hse.ru

Е.Б. Кротова, преподаватель, ekrotova@hse.ru

Д.А. Скоринкин, преподаватель, dskorinkin@hse.ru

А.С. Шишкова, преподаватель, ashishkova@hse.ru

Одобрена на заседании департамента больших данных и информационного поиска
«__» _____ 2018 г.

Руководитель департамента В.В. Подольский
_____ [подпись]

Утверждена Академическим советом образовательной программы
«__» _____ 2018 г., № протокола _____

Академические руководители образовательных программ «Языки и литература Индии», «Язык и
литература Ирана», «Библеистика и история древнего Израиля»

Комиссаров Д.А. _____ [подпись]

Арманд Е.Е. _____ [подпись]

Селезнёв М.Г. _____ [подпись]



Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения подразделения-разработчика программы.

1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и результатам обучения студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину «Цифровая грамотность», учебных ассистентов и студентов направления подготовки/специальности 58.03.01 «Востоковедение и африканистика», обучающихся по образовательным программам «Языки и литература Индии», «Язык и литература Ирана», «Библеистика и история древнего Израиля».

Программа учебной дисциплины разработана в соответствии с:

- ФГОС ВО/Образовательным стандартом НИУ ВШЭ по направлению подготовки 58.03.01 «Востоковедение и африканистика»;
- Образовательными программами «Языки и литература Индии», «Язык и литература Ирана», «Библеистика и история древнего Израиля».

2 Цели освоения дисциплины

Целями освоения дисциплины «Цифровая грамотность» являются:

- Ознакомление студентов с основами науки о данных и принципами работы современных технологий
- Ознакомление студентов с задачами Digital Humanities, специфическими для их профильной дисциплины.
- Формирование у студентов практических навыков работы с данными и решения прикладных задач анализа и визуализации данных.
- Формирование у студентов практических навыков работы с корпусными приложениями, библиотечными менеджерами и офисными технологиями.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

Уровни формирования компетенций:

РБ — ресурсная база, в основном теоретические и предметные основы (знания, умения);

СД – способы деятельности, составляющие практическое ядро данной компетенции;

МЦ – мотивационно-ценностная составляющая, отражает степень осознания ценности компетенции человеком и готовность ее использовать

В результате освоения дисциплины студент осваивает компетенции:

Компетенция	Код по ОС ВШЭ	Уровень формирования компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Способен учиться, приобретать новые	УК-1	РБ	Студент воспринимает и осваивает информацию,	Практические занятия.



знания, умения, в том числе в области, отличной от профессиональной			изложенную на практических занятиях и в результате качественно работает на практических занятиях по пройденным темам.	
Способен работать с информацией: находить, оценивать и использовать информацию из различных источников, необходимую для решения научных и профессиональных задач (в том числе на основе системного подхода)	УК-5	РБ	Студент дополняет свои знания материалами из общедоступных источников и в результате как следует выполняет практические домашние работы (проекты).	Практические домашние работы (проекты).

4 Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к циклу дисциплин Общего цикла, код Б.О. Так как ее изучение начинается на первом курсе, изначально она базируется на предметных и метапредметных умениях, приобретенных в результате освоения ФГОС Среднего полного (общего) образования и не требует пререквизитов.

Навыки использования цифровых технологий для решения различного типа учебных и профессиональных задач, сформированные дисциплиной Цифровая грамотность, могут быть использованы при изучении следующих дисциплин профессионального цикла: Межкультурная деловая коммуникация, Теория и практика устного и письменного перевода, Педагогическая психология и методика преподавания иностранных языков, а также в проектной деятельности и проведении исследований в рамках научно-исследовательского семинара, выполнения курсовой и выпускной квалификационной работы.

5 Тематический план учебной дисциплины

№	Название раздела	Всего часов	Аудиторные часы	Самостоятельная работа
			Практические занятия	
1	Данные в науке и жизни, поиск научной литературы, сбор и визуализация данных, работа в текстовых редакторах.	57	10	47
2	Корпусная лингвистика и корпусные технологии, основы	57	10	47



	статистики и научный метод, параллельные корпуса.			
	ИТОГО	114	20	94

6 Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год				Параметры
		1	2	3	4	
Текущий	Практические домашние работы (проекты)				3	Исследование по тематике семинаров, результатом которого является отчет или выступление (защита)
Промежуточный	Контрольная работа				1	Письменная работа на 40 минут
Итоговый	Экзамен				1	Выставляется по накопленной оценке

7 Критерии оценки знаний, навыков

В курсе предусмотрено несколько форм контроля знания:

- Практические домашние работы (проекты)
- Контрольная работа

Оценки по всем формам контроля выставляются по 10-ти балльной шкале.

При прохождении контроля студент должен продемонстрировать понимание основных типов и источников данных, умение проводить аналитику на данных (вычисление статистик, визуализация), использовать корпуса для сбора текстовых данных, проводить первичный анализ данных, решать задачи, не требующие сложной подготовки данных.

8 Содержание дисциплины

1. Данные в науке и жизни, поиск научной литературы, сбор и визуализация данных, работа в текстовых редакторах (5 семинаров).

Данные в науке и жизни, Поиск научной литературы, Расширенный поиск в поисковых системах, Работа в текстовых редакторах, Библиографические менеджеры, Онлайн-редакторы, Сбор и визуализация данных, Выполнение и защита проекта по теме.

2. Корпусная лингвистика и корпусные технологии, основы статистики и научный метод, параллельные корпуса (5 семинаров).

Корпусные технологии, Работа с корпусами и корпусными менеджерами, Национальный корпус русского языка, Частотность, Лексическое разнообразие, Облака тэгов, Параллельные корпуса, Лингвоспецифичная лексика, Выполнение проекта по теме и написание отчета.



9 Оценочные средства для текущего контроля и аттестации студента

Типовые фрагменты заданий домашней работы:

- Сделать графическую визуализацию данных НКРЯ
- Проанализировать корпус с помощью инструмента AntConc
- Построить диаграммы и графики, позволяющие визуализировать результаты исследования
- Подготовить отчет, используя средства форматирования текста

Типовые фрагменты заданий контрольной работы:

- В чем функция библиографических менеджеров? Какие между ними есть отличия?
- Какие библиографические и реферативные базы данных Вы знаете? Чем они отличаются друг от друга?
- Что называют стоп-листом в корпусной лингвистике? Каким образом его получают?
- Какое исследование можно было бы провести с помощью параллельного корпуса?

10 Порядок формирования оценок по дисциплине

Подведение итогов аттестации по дисциплине проводится только на основании результатов текущего контроля без проведения аттестационного испытания в форме экзамена, т.е.:

$$O_{\text{итог}} = O_{\text{накоп}}$$

Накопленная оценка рассчитывается по формуле:

$$O_{\text{накоп}} = 0.2 \cdot O_{\text{проект1}} + 0.3 \cdot O_{\text{проект2}} + 0.3 \cdot O_{\text{проект3}} + 0.2 \cdot O_{\text{кр}}$$

где $O_{\text{проект1}}$, $O_{\text{проект2}}$, $O_{\text{проект3}}$ — оценки за практические домашние работы (проекты) 1, 2 и 3 соответственно,

$O_{\text{кр}}$ — оценка за контрольную работу.

Результирующая оценка по учебной дисциплине округляется к ближайшему целому.

11 Учебно-методическое и информационное обеспечение дисциплины

11.1 Базовые учебники

- Б.Г. Миркин (2014). Введение в анализ данных: учебник и практикум для бакалавриата и магистратуры. М.: Юрайт.

11.2 Дополнительные источники

- Schreibman S., Siemens R., Unsworth, J. (eds.) (2004) Companion to Digital Humanities. Blackwell Companions to Literature and Culture. Oxford: Blackwell.
- Michel, J.B., (2011). Quantitative analysis of culture using millions of digitized books . Science, 331(6014): pp. 176–82.

Полезные Интернет-ресурсы:

- <https://digitalhumanities.umass.edu/projects>



- <http://dirtdirectory.org/>
- <http://digitalhumanities.berkeley.edu/resources/text-analysis-resources>
- <http://www.ruscorpora.ru/>
- <https://www.sketchengine.co.uk/>
- <https://programminghistorian.org/lessons/corpus-analysis-with-antconc>
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4551210/>
- <https://tone-analyzer-demo.ng.bluemix.net/>