

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 27.11.2023 18:33:02

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

Автономная некоммерческая образовательная организация высшего образования  
«Европейский университет в Санкт-Петербурге»

УТВЕРЖДАЮ

Ректор

Волков В.В.

« 31 » августа 2021 г.

Протокол Ученого Совета

№ 7 от 31 августа 2021 г.



Рабочая программа дисциплины  
«Разметка данных»

дополнительная профессиональная программа  
«Прикладной анализ данных»

вид программы  
программа профессиональной переподготовки

язык обучения – русский  
форма обучения – очная

Санкт-Петербург

**Авторы:**

Браславский П.И, кандидат технических наук, доцент факультета социологии и философии АНООВО «ЕУСПб»

Рабочая программа дисциплины «Разметка данных», входящая в состав дополнительной профессиональной программы профессиональной переподготовки «Прикладной анализ данных» утверждена на заседании Ученого совета университета.

**Содержание**

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ.....	4
2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ.....	4
3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ.....	5
4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ .....	6
5. ОЦЕНОЧНЫЕ СРЕДСТВА.....	7
6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	9
7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ.....	11

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ

**Цель обучения:** освоения дисциплины «Разметка данных» — получение навыков работы с краудсорсинговой платформой Я.Толока и встраивания краудсорсинга в проекты с машинным обучением.

### **Задачи обучения:**

- Знакомство с концепциями краудсорсинга, его роли в бизнес-процессах и исследованиях.
- Получение навыков декомпозиции исследовательской задачи и превращение ее в задачу разметки.
- Получение навыков составления инструкций для краудсорсингового задания.
- Знание основных приемов контроля качества ассессоров.
- Получение навыков работы с сервисом сервиса «ТОЛОКА», знакомство с основными сценариями работы.

Изучение данной дисциплины способствует формированию профессиональных навыков по разработке теоретических моделей исследуемых процессов, явлений и объектов, относящихся к сфере профессиональной деятельности, их исследованию, оценке и интерпретации полученных результатов.

Отличительной особенностью реализуемого подхода к преподаванию дисциплины является разнообразных практических иллюстраций основных теоретических положений.

Общая трудоемкость дисциплины составляет 54 часа.

## 2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины слушатель должен приобрести следующие знания и умения, необходимые для качественного изменения и (или) получения новых профессиональных компетенций:

### **слушатель должен знать:**

- основные концепции краудсорсинга;
- роль краудсорсинга в бизнес-процессах;
- основные принципы составления инструкций для краудсорсингового задания;
- знание приемов контроля качества ассессоров;
- основные сценарии работы с сервисом «ТОЛОКА».

### **слушатель должен уметь:**

- использовать знание приемов контроля качества ассессоров;
- правильно оформлять и представлять результаты работы.

### **слушатель должен владеть:**

- навыками декомпозиции исследовательской задачи и превращение ее в задачу разметки;
- навыками работы с сервисом «ТОЛОКА»;
- способностью к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности.

В результате изучения дисциплины «Разметка данных» слушатель приобретает следующие профессиональные компетенции (Таблица 1):

## Планируемые результаты обучения по дисциплине

Таблица 1

Код и название компетенции	Содержание компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапам формирования компетенций
ПК-2	способен организовать аналитическую работу в IT-проекте	<b>Знать:</b> <b>З (ПК-2)</b> – теоретические основы организации аналитической работы в IT-проекте
		<b>Уметь:</b> <b>У (ПК-2)</b> – анализировать данные необходимые для аналитической работы и распределять роли в IT-проекте
		<b>Владеть:</b> <b>В (ПК-2)</b> – навыками организации аналитической работы в IT-проекте

### 3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) в соответствии с таблицей 2.

Таблица 2

#### Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Коды ЗУВ (в соответствии с табл. 1)
1	Введение и общая концепция краудсорсинга	Краудсорсинг как особый способ построение бизнес-процессов. Платформа Толока. Ключевые этапы работы. Декомпозиция задачи.	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
2	Декомпозиция исходной задачи	Декомпозиция исходной задачи: как перевести исследовательский дизайн в задачу разметки. Отбор исполнителей. Популярные модели агрегации	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
3	Инструкция для краудсорсингового задания	Почему инструкция важна. Типовая структура задания. Форматирование. Иллюстрации.	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
4	Контроль качества ассессоров: inter-rater reliability et al. (+ ЕУ)	Контроль качества ассессоров: inter-rater reliability et al. (+ ЕУ)	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
5	Динамическое перекрытие и прайсинг	Пул заданий. Механизм динамического перекрытия. Подготовка полей выходных данных.	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
6	Сценарий «Опрос»	Проблема смещенности ассессоров — можно ли генерализовать результаты.	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
7	Сценарий «Расшифровка аудио»	Специфика сценария. Шаблон "Транскрипция аудио". Поле входных данных "аудио".	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
8	Практикум: разработка дизайна проекта, использующего «Толоку»		ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)

## Структура дисциплины

Таблица 3.

№ п/п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины, час.				Форма текущего контроля успеваемости, промежуточной аттестации
			Всего	Аудиторная работа по видам учебных занятий		СР <sup>1</sup>	
				Л	СЗ <sup>2</sup>		
1.	Введение и общая концепция краудсорсинга	Краудсорсинг как особый способ построение бизнес-процессов. Платформа Толока. Ключевые этапы работы. Декомпозиция задачи.	8	2	2	4	Практические задания, опросы, диспуты
2.	Декомпозиция исходной задачи	Декомпозиция исходной задачи: как перевести исследовательский дизайн в задачу разметки. Отбор исполнителей. Популярные модели агрегации	8	2	2	4	
3.	Инструкция для краудсорсингового задания	Почему инструкция важна. Типовая структура задания. Форматирование. Иллюстрации.	8	2	2	4	
4.	Контроль качества ассессоров: inter-rater reliability et al. (+ ЕУ)	Контроль качества ассессоров: inter-rater reliability et al. (+ ЕУ)	9	2	2	5	
5.	Динамическое перекрытие и прайсинг	Пул заданий. Механизм динамического перекрытия. Подготовка полей выходных данных.	8	2	2	4	
6.	Сценарий «Опрос»	Проблема смещенности ассессоров — можно ли генерализовать результаты.	4	1	1	2	
7.	Сценарий «Расшифровка аудио»	Специфика сценария. Шаблон "Транскрипция аудио". Поле входных данных "аудио".	4	1	1	2	
8.	Практикум: разработка дизайна проекта, использующего «Толоку»		8	2	2	4	
9.	Промежуточная аттестация	Опрос по вопросам (устно)	2	-	-	2	
Всего:			54	14	14	26	

## 4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

### Общие положения.

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения исследовательской литературы (из списков основной, дополнительной), статей по проблематике занятия и их анализа.

Самостоятельная работа обучающегося представляет самостоятельное изучение дополнительных материалов, Интернет-ресурсов и пр. Подготовка к семинарским занятиям, выполнение практических заданий, создание докладов, проектов и презентаций также является важной формой работы обучающихся. Самостоятельная работа может вестись как

<sup>1</sup> Самостоятельная работа, включает в себя часы на текущий контроль

<sup>2</sup> Могут включать в себя: лабораторные работы, круглые столы, мастер-классы, мастерские, деловые игры, ролевые игры, тренинги, семинары по обмену опытом, выездные занятия, консультации

индивидуально, так и при содействии преподавателя. Вопросы и замечания, возникшие в ходе самостоятельного внеаудиторного чтения рекомендованной литературы, обсуждаются с преподавателем и другими обучающимися. Выносятся на обсуждение, как правило, актуальные проблемы и предлагается их рассмотреть с точки зрения того или иного теоретического подхода.

На занятиях материал излагается в проблемной форме. Основной упор в преподавании делается на изучение теоретических понятий и возможности их применения на конкретных примерах, в том числе в устных выступлениях обучающихся.

#### **Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся.**

- Какова роль краудсорсинга в бизнес-процессах и исследованиях?
- Какие площадки для краудсорсинга используются? Какими требованиям должны удовлетворять такие площадки?
- Каковы основные этапы составления инструкции для ассессоров?
- Какие основные сценарии работы с сервисом «Толока»?
- Какие действия с заданиями можно автоматизировать в «Толоке»? Каковы возможности ее API?
- Как проверяется качество ассессоров?
- Как получать и обрабатывать данные, собранные ассессорами в «Толоке»?

#### **Источники для самостоятельной подготовки:**

- Аньшин, В.М. Управление проектами: фундаментальный курс / В.М. Аньшин, А.В. Алешин, К.А. Багратиони; ред. В.М. Аньшин, О.М. Ильина. – Москва: Издательский дом Высшей школы экономики, 2013. – 624 с. – (Учебники Высшей школы экономики). – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=227270> (
- Вылегжанина, А.О. Информационно-технологическое и программное обеспечение управления проектом: учебное пособие / А.О. Вылегжанина. – Москва; Берлин: Директ-Медиа, 2015. – 429 с.: ил., схем., табл. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=362892>
- Гринберг, А.С. Информационный менеджер: учебное пособие / А.С. Гринберг, И.А. Король. – Москва: Юнити, 2015. – 415 с. – (Профессиональный учебник: Информатика). – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=114421>
- Каширина, А.М. Развитие информационного общества: учебное пособие: [16+] / А.М. Каширина; Новосибирский государственный технический университет. – Новосибирск: Новосибирский государственный технический университет, 2019. – 92 с: ил., табл. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=576339>

## **5. ОЦЕНОЧНЫЕ СРЕДСТВА**

Проведение текущего контроля в рамках реализации данной дисциплины проходит в соответствии с Таблицей 3 данной рабочей программы дисциплины по основным понятиям (категориям) и проблемам, рассматриваемым в предложенных темах. Фиксация результатов текущего контроля в рамках реализации данной дисциплины не предусмотрена.

#### **Типовые задания к текущей аттестации (практические задания, опросы, диспуты).**

**Опрос 1.** Слушателям предлагается привести не менее трех обоснованных тезисов по проблеме:

1. Какова общая концепция краудсорсинга?
2. Охарактеризуйте краудсорсинг как особый способ построение бизнес-процессов.
3. Опишите основные принципы работы платформы Толока.
4. Каковы ключевые этапы работы краудсорсинга?

5. Каковы методы декомпозиции задачи?
6. Сделать обзор крупных краудсорсинговых проектов (Википедия, Zooniverse и т.д.) и информационных технологий, на которых они построены.

Практическое задание 1. Слушателям предлагается реализовать в Толоке сформулированную на практике задачу разметки.

Практическое задание 2. Слушатели получают инструкцию для краудсорсингового задания. Задача слушателя состоит в том, чтобы исправить выданную инструкцию, аргументировать внесенные изменения.

Опрос 2. Слушателям предлагается привести не менее трех обоснованных тезисов по проблеме:

2. Слушателям предлагается проанализировать качество ассессоров в одном из проектов.
3. Слушателям необходимо сделать выводы о процедурах и технических средствах такого контроля.

Практическое задание 3. Слушателям необходимо сравнить выполнение одного проекта с/без динамического перекрытия и управления прайсингом.

Практическое задание 4. Слушателям предлагается получить и обработать результаты задачи, сделанной на практике.

Практическое задание 5, диспут 1. Слушателям необходимо получить и обработать результаты задачи, сделанной на практике.

### Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Опрос	<p>Ответ отсутствует или является односложным, или содержит существенные ошибки – не зачтено</p> <p>Слушатель в ответах демонстрирует знание всех теоретических положений, (развернуто) отвечает на все поставленные вопросы, предлагает обоснования при ответе на все или большинство поставленных вопросов; несущественные ошибки не снижают качество ответа — зачтено</p>
Диспут	<p>Пассивность, участие без представления аргументов и обоснования точки зрения, несформированность навыков профессиональной коммуникации в группе — не зачтено</p> <p>Представление аргументированной научной позиции, обоснование точки зрения в диспуте, демонстрация навыков профессиональной коммуникации в группе — зачтено</p>
Практическое задание	<p>Слушатель выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено,</p> <p>полное и правильное выполнение задания в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено</p>

Форма промежуточной аттестации – зачет, выставляемый на основе проекта.

При аттестации используются система «зачтено» и «не зачтено» в соответствии с критериями оценивания.

В результате промежуточного контроля знаний обучающиеся получают аттестацию по дисциплине.

**Показатели, критерии и оценивание компетенций по уровням их формирования в процессе промежуточной аттестации**

Таблица 4

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Коды ЗУВ (в соответствии с Таблицей 1)	Критерии оценивания	Оценка
зачет / проект	ПК-4	З (ПК-4) У (ПК-4) В (ПК-4)	<p>Слушатель демонстрирует полную самостоятельность в подборе фактического материала и аналитическое отношение к нему, умение рассматривать примеры и факты во взаимосвязи и взаимообусловленности, отбирать наиболее существенные из них; а также показывает грамотное использование методов описания и презентации исследования</p> <p>Слушатель демонстрирует полную самостоятельность в подборе фактического материала и аналитическое отношение к нему, умение рассматривать примеры и факты во взаимосвязи и взаимообусловленности, отбирать наиболее существенные из них; а также показывает грамотное использование методов описания и презентации исследования</p>	зачтено
			<p>Слушатель не демонстрирует аналитическое отношение к материалу, не видит взаимосвязь примеров и фактов; а также использует методы описания и презентации исследования с большим количеством существенных ошибок</p>	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», показывают уровень сформированности у обучающегося компетенций.

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «не зачтено», показывают не сформированность у обучающегося компетенций по дисциплине.

**Типовые темы проектов к промежуточной аттестации.**

- Запустить крауд-тестирование в Толоке по исследованию интерфейса и функционала отдельно взятого поискового Колдунщика Яндекса. Поискового Колдунщика каждый слушатель выбирает самостоятельно.
- Постановка цели: поиск ошибок/пользовательских проблем, о которых не знают продуктовые команды.
- Написать тестовые сценарии для исследования, выбранного Колдунщика.
- Создать Проект в Толоке для тестирования, доработать в нём интерфейс.
- Подумать над контролем качества/мотивацией/обработкой результата.
- Раздать тестовые задания на толokersов.
- Обработать полученные результаты.

**6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

**Основная литература:**

- Аньшин, В.М. Управление проектами: фундаментальный курс / В.М. Аньшин, А.В. Алешин, К.А. Багратиони; ред. В.М. Аньшин, О.М. Ильина. – Москва: Издательский дом Высшей школы экономики, 2013. – 624 с. – (Учебники Высшей школы экономики). – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=227270> (
- Вылегжанина, А.О. Информационно-технологическое и программное обеспечение управления проектом: учебное пособие / А.О. Вылегжанина. – Москва; Берлин: Директ-Медиа, 2015. – 429 с: ил., схем., табл. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=362892>

#### **Дополнительная литература:**

- Гринберг, А.С. Информационный менеджмент: учебное пособие / А.С. Гринберг, И.А. Король. – Москва: Юнити, 2015. – 415 с. – (Профессиональный учебник: Информатика). – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=114421>
- Каширина, А.М. Развитие информационного общества: учебное пособие: [16+] / А.М. Каширина; Новосибирский государственный технический университет. – Новосибирск: Новосибирский государственный технический университет, 2019. – 92 с: ил., табл. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=576339>.

Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения программы:

#### **Информационно-справочные системы:**

- Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
- Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
- Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
- Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
- Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
- Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
- Российское образование. Федеральный портал: <http://www.edu.ru>

#### **Тематические системы:**

- Google. Книги: <https://books.google.com>
- Internet Archive: <https://archive.org>
- Кооб.ру. Электронная библиотека «Куб»: <http://www.koob.ru/philosophy/>
- Библиотека Ихтика [ihtik.lib.ru]: <http://ihtik.lib.ru/>
- Докусфера — Российская национальная библиотека: <http://leb.nlr.ru>
- ЕНИП — Электронная библиотека «Научное наследие России»: <http://e-heritage.ru/>
- Интелрос. Интеллектуальная Россия: <http://www.intelros.ru/>
- Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
- Неприкосновенный запас: <http://magazines.russ.ru/nz/>
- Президентская библиотека: <http://www.prlib.ru>
- Российская государственная библиотека: <http://www.rsl.ru/>
- Российская национальная библиотека: <http://www.nlr.ru/poisk/>

## **7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ**

В ходе реализации образовательного процесса используются многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа и семинарского типа обеспечивается демонстрационным оборудованием.

Для обучающихся с ограниченными возможностями здоровья и инвалидов (в случае необходимости) могут быть созданы специальные условия для получения образования.

### **Программное обеспечение**

При осуществлении образовательного процесса в рамках Университета слушателям рекомендовано использовать следующее лицензионное программное обеспечение:

- OS Microsoft Windows (OVS OS Platform)
- MS Office (OVS Office Platform)
- Adobe Acrobat Professional 11.0 MLP AOO License RU
- Adobe CS5.5 Design Standart Win IE EDU CLP
- ABBYY FineReader 11 Corporate Edition
- ABBYY Lingvo x5
- Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
- Adobe Acrobat Reader DC /Pro – бесплатно
- Google Chrome – бесплатно
- Opera – бесплатно
- Mozilla – бесплатно
- VLC – бесплатно