

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 06.10.2026 11:32:04

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

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

УТВЕРЖДАЮ



Ректор /В.В. Волков/

«15» октября 2026 г.

Протокол Ученого совета

№ 9 от 30 сентября 2026 г.

**Рабочая программа дисциплины
«АНАЛИТИКА ДАННЫХ И АВТОМАТИЧЕСКОЕ ИЗВЛЕЧЕНИЕ СКРЫТЫХ
СВЯЗЕЙ»**

Дополнительная профессиональная программа
«ЦИФРОВЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ КУЛЬТУРНОГО НАСЛЕДИЯ»

вид программы
программа повышения квалификации

язык обучения – русский
форма обучения – очная (очная с применением дистанционных образовательных технологий)

Санкт-Петербург

Автор:

Старобыховская А.А., руководитель направления компьютерного зрения и разработки нейросетей, Лаборатория «Искусство и искусственный интеллект» АНООВО «ЕУСПб».

Рабочая программа дисциплины «Аналитика данных и автоматическое извлечение скрытых связей», входящая в состав дополнительной профессиональной программы повышения квалификации «Цифровые методы исследования культурного наследия», утверждена на заседании Ученого совета университета.

Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ.....	4
2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ	4
2. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ	6
4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ.....	9
5. ОЦЕНОЧНЫЕ СРЕДСТВА	10
6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	12
7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ДИСЦИПЛИНЫ	13

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ

Цель обучения по дисциплине «Аналитика данных и автоматическое извлечение скрытых связей» состоит в формировании навыков использования цифровых методов для поиска информации, выявления взаимосвязей и аналитической обработки исследовательских данных. В рамках дисциплины рассматриваются задачи анализа визуальных и текстовых данных, возможности и ограничения машинного обучения, компьютерного зрения и искусственного интеллекта, а также способы представления найденных связей.

Задачи обучения:

- сформировать навыки поиска, классификации данных и выявления взаимосвязей между объектами исследования;
- освоить основные подходы к аналитической обработке текстовых и визуальных данных;
- сформировать навыки представления связанных данных и формулирования выводов на основе результатов их анализа.

Изучение данной дисциплины способствует формированию профессиональных навыков анализа исследуемых процессов, явлений и объектов, относящихся к сфере профессиональной деятельности, оценки и интерпретации полученных результатов. Отличительной особенностью реализуемого подхода к преподаванию дисциплины является разнообразие практических иллюстраций основных теоретических положений применительно к изучаемой сфере, что дает возможность обучающимся увязать теоретические и практические аспекты.

Общая трудоемкость дисциплины составляет 68 часов (включая самостоятельную работу и часы на промежуточную аттестацию).

2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины слушатель должен приобрести следующие знания и умения, необходимые для качественного изменения профессиональных компетенций:

слушатель должен знать:

- ☐ подходы к анализу текстовых и визуальных данных;
- ☐ способы выявления взаимосвязей между объектами исследования;
- ☐ возможности и границы применения цифровых инструментов, машинного обучения, компьютерного зрения и искусственного интеллекта;

- ☐ основные способы представления связанных данных;
слушатель должен уметь:
- ☐ выявлять связи между объектами исследования и формулировать выводы на основе полученных данных;
- ☐ выбирать подходящий способ представления результатов анализа;
слушатель должен владеть:
- ☐ методами поиска, классификации и систематизации данных, а также выявления взаимосвязей между объектами исследования;
- ☐ цифровыми инструментами анализа визуальных и текстовых данных, включая классификацию, кластеризацию, распознавание объектов, автоматическое тегирование, суммаризацию и выделение именованных сущностей;
- ☐ способами представления связанных данных с использованием таблиц, графов и диалоговых систем.

В результате изучения дисциплины «Аналитика данных и автоматическое извлечение скрытых связей» слушатель приобретает следующие профессиональные компетенции (Таблица 1):

Планируемые результаты обучения по дисциплине

Таблица 1

Код и название компетенции	Содержание компетенции	Планируемые результаты обучения по программе, характеризующие этапы формирования компетенций
ПК-1	способен самостоятельно анализировать и интерпретировать данные современных научных исследований в отдельной предметной области, критически переосмысливать накопленный опыт, изменять при необходимости вид и характер своей профессиональной деятельности	Знать: З (ПК-1) - основные методы моделирования, анализа и критической оценки полученных результатов в своей предметной области
		Уметь: У (ПК-1) – осуществлять критический анализ научной литературы, предлагать интерпретацию научных данных в отдельной предметной области; использовать методологию описания процессов и явлений в сфере профессиональной деятельности; выбирать оптимальные методы в процессе решения профессиональных задач
		Владеть: В (ПК-1) – навыками систематизации и использования информации, необходимой для решения задач профессиональной деятельности; навыками создания стандартных теоретических моделей, анализа и интерпретации полученных результатов

Код и название компетенции	Содержание компетенции	Планируемые результаты обучения по программе, характеризующие этапы формирования компетенций
ПК-2	способен организовывать и осуществлять самостоятельную профессиональную деятельность с использованием цифровых инструментов работы с данными на всех этапах их жизненного цикла: от оцифровки и описания до организации хранения, поиска взаимосвязей и извлечения новых знаний	Знать: З (ПК-2) – подходы к анализу текстовых и визуальных данных; способы выявления взаимосвязей между объектами исследования; возможности и границы применения цифровых инструментов, машинного обучения, компьютерного зрения и искусственного интеллекта; основные способы представления связанных данных Уметь: У (ПК-2) – выявлять связи между объектами исследования и формулировать выводы на основе полученных данных; выбирать подходящий способ представления результатов анализа Владеть: В (ПК-2) – методами поиска, классификации и систематизации данных, а также выявления взаимосвязей между объектами исследования; цифровыми инструментами анализа визуальных и текстовых данных, включая классификацию, кластеризацию, распознавание объектов, автоматическое тегирование, суммаризацию и выделение именованных сущностей; способами представления связанных данных с использованием таблиц, графов и диалоговых систем

2. 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.	Поиск данных и выявление взаимосвязей	Введение в исследовательскую работу с данными. Формулирование исследовательских вопросов. Методы поиска информации, поиск по критериям и поиск аналогий. Выявление связей между объектами и	ПК-1	З (ПК-1) У (ПК-1) В (ПК-1)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Коды ЗУВ (в соответствии с Табл. 1)
		формирование исследовательских гипотез		
2.	Машинное обучение, компьютерное зрение и искусственный интеллект в исследовательской деятельности	Общее представление о машинном обучении, компьютерном зрении и искусственном интеллекте. Возможности их применения к исследовательским данным и границы применимости соответствующих подходов	ПК-1	З (ПК-1) У (ПК-1) В (ПК-1)
3.	Анализ визуальных данных	Информация, содержащаяся в визуальных данных, и основные задачи компьютерного зрения. Классификация и кластеризация изображений. Распознавание объектов: локализация и классификация. Генерация описаний и автоматическое тегирование изображений	ПК-1	З (ПК-1) У (ПК-1) В (ПК-1)
4.	Анализ текстовых данных	Применение цифровых подходов к работе с текстовыми массивами. Суммаризация текстов и поиск именованных сущностей	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
5.	Работа с данными различных модальностей: звук, 3D и мультимодальные системы	Ознакомление с особенностями представления и использования аудиоданных, 3D-данных и других модальностей в исследовательской деятельности. Рассматриваются общие возможности анализа таких данных и подходы к совместной работе с несколькими типами информации в мультимодальных системах	ПК-2	З (ПК-2) У (ПК-2) В (ПК-2)
6.	Отображение связанных данных	Представление результатов анализа и выявленных взаимосвязей. Использование таблиц и графов	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 промежу- точной аттестации/ итоговой аттестации
			Всего	Аудиторная работа по видам учебных занятий		СР	
				Л	СЗ ¹		
1.	Поиск данных и выявление взаимосвязей	Введение в исследовательскую работу с данными. Формулирование исследовательских вопросов. Методы поиска информации, поиск по критериям и поиск аналогий. Выявление связей между объектами и формирование исследовательских гипотез	9	1	3	5	Устный опрос
2.	Машинное обучение, компьютерное зрение и искусственный интеллект в исследовательской деятельности	Общее представление о машинном обучении, компьютерном зрении и искусственном интеллекте. Возможности их применения к исследовательским данным и границы применимости соответствующих подходов	11	3	3	5	
3.	Анализ визуальных данных	Информация, содержащаяся в визуальных данных, и основные задачи компьютерного зрения. Классификация и кластеризация изображений. Распознавание объектов: локализация и классификация. Генерация описаний и автоматическое тегирование изображений	11	3	3	5	
4.	Анализ текстовых данных	Применение цифровых подходов к работе с текстовыми массивами. Суммаризация текстов и	11	3	3	5	

¹ СЗ – семинарские занятия могут включать в себя: лабораторные работы, круглые столы, мастер-классы, мастерские, деловые игры, ролевые игры, тренинги, семинары по обмену опытом, выездные занятия, консультации.

№ п/п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины час.			Форма промежу- точной аттестации/ итоговой аттестации	
			Всего	Аудиторная работа по видам учебных занятий			СР
				Л	СЗ ¹		
		поиск именованных сущностей					
5.	Работа с данными различных модальностей: звук, 3D и мультимодальные системы	Ознакомление с особенностями представления и использования аудиоданных, 3D-данных и других модальностей в исследовательской деятельности. Рассматриваются общие возможности анализа таких данных и подходы к совместной работе с несколькими типами информации в мультимодальных системах	12	3	4	5	
6.	Отображение связанных данных	Представление результатов анализа и выявленных взаимосвязей. Использование таблиц и графов	12	3	4	5	
7.	Промежуточная аттестация		2			2	
Итого по дисциплине:			68	16	20	32	

4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

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 и дополнительной литературы), статей по проблематике занятия и их анализа.

Самостоятельная работа обучающегося включает изучение дополнительных материалов, интернет-ресурсов и других источников информации. Самостоятельная работа может вестись как 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.

Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся.

– Аналитика и извлечение связей — исследовательский вопрос, методы поиска, выявление взаимосвязей между объектами, построение гипотез.

– Цифровые инструменты и искусственный интеллект — машинное обучение, компьютерное зрение, работа с текстовыми и визуальными данными, суммаризация, автоматическое тегирование.

– Визуализация данных — формы представления связанных данных: таблицы, графы, чат-системы, способы наглядного отображения результатов анализа.

Источники для самостоятельной подготовки:

– Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Юрайт, 2014.

– Парад цифровых гуманитарных проектов : монография / А. Е. Арёфьев, Т. А. Базарова, Е. В. Баранова [и др.] ; отв. ред. А. Ю. Володин. — Красноярск : Сибирский федеральный университет, 2025.

– Костыгов, Н. Д. Опыт и перспективы применения искусственного интеллекта для исследования иероглифических рукописей на вымерших языках / Н. Д. Костыгов, Е. В. Ягунова // Культура и технологии. — 2026. — Т. 11, вып. 1. — С. 17–28. — DOI 10.17586/2587-800X-2026-11-1-17-28.

– Юмашева, Ю. Ю. К вопросу о применении искусственного интеллекта в исторических исследованиях // Историческая информатика. — 2025. — № 1. — С. 95–121. — DOI 10.7256/2585-7797.2025.1.73578.

– Унтерберг, А. М. Распознавание рукописного текста исторических документов с применением технологий глубоких нейронных сетей / А. М. Унтерберг, А. В. Пятаева, С. С. Замыслова [и др.] // Известия высших учебных заведений. Приборостроение. — 2024. — Т. 67, № 9. — С. 767–775. — DOI 10.17586/0021-3454-2024-67-9-767-775.

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.

Форма промежуточной аттестации – зачет, выставляемый на основе устного опроса.

При аттестации используется бинарная система («зачтено» / «не зачтено») в соответствии с критериями оценивания.

В результате промежуточного контроля знаний обучающиеся получают аттестацию по дисциплине.

Перечень вопросов к зачету по дисциплине приведен в Оценочных материалах для проведения промежуточной аттестации по программе.

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
Устный опрос / Зачет	ПК-1 ПК-2	З (ПК-1) У (ПК-1) В (ПК-1) З (ПК-2) У (ПК-2) В (ПК-2)	Ответ слушателя показывает: - глубокое знание содержания дисциплины; - последовательное, логически выстроенное изложение; - владение теоретическими положениями и их грамотная аргументация; - соблюдение норм устной литературной речи	Зачтено
			Ответ слушателя показывает: - незнание значительной части программного материала, - наличие существенных ошибок в определениях, формулировках, понимании теоретических положений; - бессистемность знания дисциплины, - отсутствие логически корректного анализа, аргументации, классификации; - наличие нарушений норм устной литературной речи	Не зачтено

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

Результаты промежуточного контроля, выраженные в бинарной системе «не зачтено», указывают на несформированность у обучающегося требуемых компетенций.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

– Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Юрайт, 2014.

Дополнительная литература:

– Парад цифровых гуманитарных проектов : монография / А. Е. Арэфьев, Т. А. Базарова, Е. В. Баранова [и др.] ; отв. ред. А. Ю. Володин. — Красноярск : Сибирский федеральный университет, 2025.

– Костыгов, Н. Д. Опыт и перспективы применения искусственного интеллекта для исследования иероглифических рукописей на вымерших языках / Н. Д. Костыгов, Е. В. Ягунова // Культура и технологии. — 2026. — Т. 11, вып. 1. — С. 17–28. — DOI 10.17586/2587-800X-2026-11-1-17-28.

– Юмашева, Ю. Ю. К вопросу о применении искусственного интеллекта в исторических исследованиях // Историческая информатика. — 2025. — № 1. — С. 95–121. — DOI 10.7256/2585-7797.2025.1.73578.

– Унтерберг, А. М. Распознавание рукописного текста исторических документов с применением технологий глубоких нейронных сетей / А. М. Унтерберг, А. В. Пятаева, С. С. Замыслова [и др.] // Известия высших учебных заведений. Приборостроение. — 2024. — Т. 67, № 9. — С. 767–775. — DOI 10.17586/0021-3454-2024-67-9-767-775.

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

Информационно-справочные системы:

- Гарант.Ру. Информационно-правовой портал: <http://www.garant.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>

Тематические системы:

- Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
- Президентская библиотека: <http://www.prilib.ru>

- Российская государственная библиотека: <http://www.rsl.ru/>
- Российская национальная библиотека: <http://www.nlr.ru/poisk/>

Профессиональные базы данных:

Полный перечень доступных обучающимся профессиональных баз данных представлен на официальном сайте Университета <https://eusp.org/library/electronicresources>, включая следующие базы данных:

- **eLIBRARY.RU** — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, наукометрическая база данных: <https://elibrary.ru>;
- Электронные журналы по подписке (текущие номера научных зарубежных журналов).

Электронные библиотечные системы:

1. Znanium.com – Электронная библиотечная система (ЭБС) – <https://znanium.ru/>;
2. Университетская библиотека онлайн – Электронная библиотечная система (ЭБС) – <https://biblioclub.ru/>

7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ДИСЦИПЛИНЫ

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.

В университете созданы все условия для функционирования электронной информационно-образовательной среды, включающей в себя электронные информационные ресурсы, электронные образовательные ресурсы, совокупность информационных технологий, телекоммуникационных технологий, соответствующих технологических средств, обеспечивающих освоение обучающимися образовательных программ в полном объеме независимо от места нахождения обучающихся. В случае реализации программы повышения квалификации с применением дистанционных образовательных технологий слушатели получают доступ ко всем необходимым ресурсам.

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:

- ABBYY FineReader 11 Corporate Edition
- Mozilla Firefox – бесплатно
- MS Office (OVS Office Platform)
- Opera – бесплатно
- OS Microsoft Windows (OVS OS Platform)
- VLC – бесплатно
- Яндекс.Браузер (Yandex Browser) – бесплатно
- МТС Линк