

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Галунин Сергей Александрович
Должность: проректор по учебной работе
Дата подписания: 18.07.2023 17:12:32
Уникальный программный ключ:
08ef34338325bdb0ac5a47baa5472ce36cc3fc3b

Приложение к ОПОП
«Системы искусственного интел-
лекта»



СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

МИНОБРНАУКИ РОССИИ

федеральное государственное автономное образовательное учреждение высшего образования
**«Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И.Ульянова (Ленина)»
(СПбГЭТУ «ЛЭТИ»)»**

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

**«ЭТИКА ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА С ИСКУССТВЕННЫМ
ИНТЕЛЛЕКТОМ»**

для подготовки бакалавров

по направлению

01.03.02 «Прикладная математика и информатика»

по профилю

«Системы искусственного интеллекта»

Санкт-Петербург

2023

ЛИСТ СОГЛАСОВАНИЯ

Разработчики:

ассистент И. А.

старший преподаватель Баженов С.С.

Рабочая программа рассмотрена и одобрена на заседании кафедры ФЛ

06.09.2022, протокол № 1

Рабочая программа рассмотрена и одобрена учебно-методической комиссией

ГФ, 09.09.2022, протокол № 3

Согласовано в ИС ИОТ

Начальник ОМОЛА Загороднюк О.В.

1 СТРУКТУРА ДИСЦИПЛИНЫ

Обеспечивающий факультет	ГФ
Обеспечивающая кафедра	ФЛ
Общая трудоемкость (ЗЕТ)	3
Курс	3
Семестр	5
Виды занятий	
Лекции (академ. часов)	17
Практические занятия (академ. часов)	17
Иная контактная работа (академ. часов)	1
Все контактные часы (академ. часов)	35
Самостоятельная работа, включая часы на контроль (академ. часов)	73
Всего (академ. часов)	108
Вид промежуточной аттестации	
Дифф. зачет (курс)	3

2 АННОТАЦИЯ ДИСЦИПЛИНЫ

«ЭТИКА ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА С ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ»

Дисциплина «Этика взаимодействия человека и систем ИИ» входит в вариативную часть общенаучного цикла подготовки бакалавров. Целью изучения дисциплины является ознакомление студентов с философскими аспектами искусственного интеллекта; усвоение базовых категорий теоретической и прикладной этики; формирование навыков применения этих категорий при анализе конкретных кейсов применительно к технологиям искусственного интеллекта.

SUBJECT SUMMARY

«ETHICS OF HUMAN AND AI RELATIONS»

The discipline «Ethics of human and AI relations» is included into a variable unit of a general scientific cycle of training of bachelor students. The purpose of studying of discipline is acquaintance of students with philosophical aspects of artificial intelligence; mastering of the basic categories of theoretical and applied ethics; developing skills in applying these categories in the framework of general expertise of artificial intelligence technologies.

3 ОБЩИЕ ПОЛОЖЕНИЯ

3.1 Цели и задачи дисциплины

1. Изучение философских аспектов применения технологий искусственного интеллекта.
2. Формирование готовности к этической оценке взаимодействие человека с системами искусственного интеллекта в изменяющихся условиях
3. В результате освоения курса обучающийся должен продемонстрировать знание этических теорий, проблему сознания и искусственного интеллекта, угрозы, связанные с внедрением ИИ.
4. Выработка умений самостоятельного использования и применения базовых категорий современной теоретической и прикладной этики в контексте комплексной этической экспертизы технологий искусственного интеллекта
5. Освоение базовых навыков принятия решений в моральных дилеммах, развитие способности самостоятельной моральной оценки применения технологий ИИ.

3.2 Место дисциплины в структуре ОПОП

Дисциплина изучается на основе ранее освоенных дисциплин учебного плана:

1. «Философия»

и обеспечивает изучение последующих дисциплин:

1. «Логический вывод в системах искусственного интеллекта»
2. «Производственная практика (преддипломная практика)»

3.3 Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения образовательной программы обучающийся должен достичь следующие результаты обучения по дисциплине:

Код компетенции/ индикатора компетенции	Наименование компетенции/индикатора компетенции
ПКО-1	Способен планировать и организовывать свою деятельность в цифровом пространстве с учетом правовых и этических норм взаимодействия человека и искусственного интеллекта и требований информационной безопасности
<i>ПКО-1.1</i>	<i>Выбирает современные технологии и системы искусственного интеллекта для решения задач в профессиональной деятельности</i>

4 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1 Содержание разделов дисциплины

4.1.1 Наименование тем и часы на все виды нагрузки

№ п/п	Наименование темы дисциплины	Лек, ач	Пр, ач	ИКР, ач	СР, ач
1	Введение	2	0		0
2	Тема 1. Основы теоретической этики	2	2		7
3	Тема 2. Современная эпистемология морали	3	3		16
4	Тема 3. Основные проблемы философии сознания	4	4	0	16
5	Тема 4. Основные подходы к определению ИИ	2	2		18
6	Тема 5. Этические проблемы применения ИИ	2	2		8
7	Тема 6. Экзистенциальные вызовы, связанные с внедрением ИИ	2	2	1	8
8	Заключение	0	2		0
	Итого, ач	17	17	1	73
	Из них ач на контроль	0	0	0	0
	Общая трудоемкость освоения, ач/зе	108/3			

4.1.2 Содержание

№ п/п	Наименование темы дисциплины	Содержание
1	Введение	Краткая аннотация учебного курса с указанием целей и основных задач курса. Введение в проблематику искусственного интеллекта и её этического осмысления.
2	Тема 1. Основы теоретической этики	1. Предмет этики. Моральное измерение личности. Моральное измерение общества. 2. Мораль в жизни человека. Этика ненасилия. Единство морали и многообразие нравов. Моральная оценка и моральное поведение. 3. Мораль и религия. Учителя человечества: Конфуций, Будда, Моисей, Иисус Христос, Мухаммед. 4. Категориальный строй этики: благо, идеал, добро и зло, долг и совесть, свобода, добродетель и порок, счастье.

№ п/п	Наименование темы дисциплины	Содержание
3	Тема 2. Современная эпистемология морали	<p>1. Основные этические парадигмы: аретология, деонтология.</p> <p>2. Основные подходы к понятию моральной ответственности. Ретрибутивизм и консеквенциализм. Принцип ответственности Х. Йонаса.</p> <p>3. Основные подходы к эпистемологии морали: социологический, психологический, онтологический, эволюционный, методологический.</p> <p>4. Моральные дилеммы. Способы решения моральных дилемм с использованием этических концепций.</p>
4	Тема 3. Основные проблемы философии сознания	<p>1. Концептуализация интеллекта в истории философской мысли. Эмпиризм и рационализм. Дуализм Р. Декарта.</p> <p>2. Основные подходы к онтологическому статусу интеллекта и разума. Редукционизм и антиредукционизм. Эпифеноменализм. Функционализм.</p> <p>3. Понятие феноменального сознания. «Аргумент Мэри» Ф. Джексона. Трудная проблема сознания Д. Чалмерса.</p> <p>4. Понятие воли. Необходимые и достаточные критерии свободы воли.</p>
5	Тема 4. Основные подходы к определению ИИ	<p>1. Тест А. Тьюринга для определения искусственного интеллекта. Понятие сильного и слабого ИИ.</p> <p>2. Аргумент китайской комнаты Дж. Серла. Основные возражения к подходу Дж. Серла.</p> <p>3. Понятие множественной реализуемости. Парадигма функционализма Х. Патнема.</p> <p>4. Моральная машина. Проблема ИИ как носителя моральной ответственности.</p>
6	Тема 5. Этические проблемы применения ИИ	<p>1. Права роботов. ИИ как объект морального действия.</p> <p>2. Место ИИ в современной правовой практике.</p> <p>3. ИИ и этическая экспертиза.</p>
7	Тема 6. Экзистенциальные вызовы, связанные с внедрением ИИ	<p>1. Сингулярность и потеря контроля над развитием ИИ. Гипотеза взрывной скорости. Пределы развития ИИ.</p> <p>2. Перспективы сосуществования ИИ и человека: уничтожение, неравенство, изоляция, искусственная эволюция человека. Мораль сверхразвитого ИИ.</p> <p>3. Развитый человеческий разум как более безопасный путь развития ИИ. Возможность расширения разума. Загрузка разума. Трансгуманизм.</p>
8	Заключение	Основные этические и экзистенциальные аспекты технологий искусственного интеллекта. Подведение итогов.

4.2 Перечень лабораторных работ

Лабораторные работы не предусмотрены.

4.3 Перечень практических занятий

Наименование практических занятий	Количество ауд. часов
1. Основы теоретической этики	2
2. Современная эпистемология морали	4
3. Основные проблемы философии сознания	4
4. Основные подходы к определению ИИ	2
5. Моральные проблемы применения технологий ИИ	2
6. Экзистенциальные вызовы, связанные с внедрением ИИ	2
7. Заключение	1
Итого	17

4.4 Курсовое проектирование

Курсовая работа (проект) не предусмотрены.

4.5 Реферат

Реферат не предусмотрен.

4.6 Индивидуальное домашнее задание

Индивидуальное домашнее задание не предусмотрено.

4.7 Доклад

Доклад не предусмотрен.

4.8 Кейс

Кейс не предусмотрен.

4.9 Организация и учебно-методическое обеспечение самостоятельной работы

Изучение дисциплины сопровождается самостоятельной работой студентов с рекомендованными преподавателем литературными источниками и ин-

формационными ресурсами сети Интернет.

Планирование времени для изучения дисциплины осуществляется на весь период обучения, предусматривая при этом регулярное повторение пройденного материала. Обучающимся, в рамках внеаудиторной самостоятельной работы, необходимо регулярно дополнять сведениями из литературных источников материал, законспектированный на лекциях. При этом на основе изучения рекомендованной литературы целесообразно составить конспект основных положений, терминов и определений, необходимых для освоения разделов учебной дисциплины.

Особое место уделяется консультированию, как одной из форм обучения и контроля самостоятельной работы. Консультирование предполагает особым образом организованное взаимодействие между преподавателем и студентами, при этом предполагается, что консультант либо знает готовое решение, которое он может предписать консультируемому, либо он владеет способами деятельности, которые указывают путь решения проблемы.

Текущая СРС	Примерная трудоемкость, ач
Работа с лекционным материалом, с учебной литературой	25
Опережающая самостоятельная работа (изучение нового материала до его изложения на занятиях)	10
Самостоятельное изучение разделов дисциплины	0
Выполнение домашних заданий, домашних контрольных работ	0
Подготовка к лабораторным работам, к практическим и семинарским занятиям	10
Подготовка к контрольным работам, коллоквиумам	12
Выполнение расчетно-графических работ	0
Выполнение курсового проекта или курсовой работы	0
Поиск, изучение и презентация информации по заданной проблеме, анализ научных публикаций по заданной теме	10
Работа над междисциплинарным проектом	0
Анализ данных по заданной теме, выполнение расчетов, составление схем и моделей, на основе собранных данных	0
Подготовка к зачету, дифференцированному зачету, экзамену	6
ИТОГО СРС	73

5 Учебно-методическое обеспечение дисциплины

5.1 Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

№ п/п	Название, библиографическое описание	К-во экз. в библ.
Основная литература		
1	Малюк А. А. Этика в сфере информационных технологий [Электронный ресурс] / А. А. Малюк, О. Ю. Полянская, И. Ю. Алексеева, 2011. -344 с.	неогр.
2	Росенко, Маргарита Николаевна. Профессиональная этика [Текст] : учеб. пособие / М.Н. Росенко, А.В. Бабаева, М.В. Чигирь ; под. общ. ред. М.Н. Росенко, 2005. -88 с	591
3	Скворцов, Алексей Алексеевич. Этика [Электронный ресурс] : Учебник и практикум для вузов / Скворцов А. А., 2021. -321 с	неогр.
Дополнительная литература		
1	Иванюк, Федор Михайлович. Этика профессиональных сообществ [Текст] : учеб. пособие / Ф.М. Иванюк, 2009. -67 с	359
2	Гусейнов, Абдусалам Абдулкеримович. Этика [Электронный ресурс] : Учебник для вузов / под общ. ред. Гусейнова А.А., 2020. -460 с	неогр.
3	Шрейдер, Юлий Анатольевич. Этика: Введение в предмет [Текст] : Учеб. пособие для вузов / Ю.А.Шрейдер, 1998. -271 с. с.	5
4	Волков, Александр Михайлович. Правовое обеспечение профессиональной деятельности в IT-сфере. Схемы, таблицы, определения, комментарии [Электронный ресурс] : Учебник Для СПО / Волков А. М., Лютягина Е. А., 2021. -281 с	неогр.

5.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», используемых при освоении дисциплины

№ п/п	Электронный адрес
1	Стэнфордская энциклопедия http://plato.stanford.edu/
2	Библиотека ИФ РАН https://iphlib.ru/library/collection/newphilenc/browse/CL6
3	Центр искусственного интеллекта НИУ ВШЭ https://cs.hse.ru/aicenter/ethics
4	Центр подготовки руководителей и команд цифровой трансформации при РАН-ХиГХ https://ethics.cdto.center/

5.3 Адрес сайта курса

Адрес сайта курса: <https://vec.etu.ru/moodle/course/view.php?id=10691>

6 Критерии оценивания и оценочные материалы

6.1 Критерии оценивания

Для дисциплины «Этика взаимодействия человека с искусственным интеллектом» предусмотрены следующие формы промежуточной аттестации: зачет с оценкой.

Зачет с оценкой

Оценка	Описание
Неудовлетворительно	Курс не освоен. Студент испытывает серьезные трудности при ответе на ключевые вопросы дисциплины
Удовлетворительно	Студент в целом овладел курсом, но некоторые разделы освоены на уровне определений и формулировок теорем
Хорошо	Студент овладел курсом, но в отдельных вопросах испытывает затруднения. Умеет решать задачи
Отлично	Студент демонстрирует полное овладение курсом, способен применять полученные знания при решении конкретных задач.

Особенности допуска

К дифференцированному зачету допускаются студенты при условии выполнения двух контрольных работ и при хорошей посещаемости практических и лекционных занятий (не менее 80 %). Итоговая оценка выставляется как среднее арифметическое по результатам текущей аттестации. Студенты, выполнившие обязательные контрольные работы, но имеющие менее 80% посещаемости, сдают дифференцированный зачет в виде собеседования по вопросам для дифференцированного зачета.

6.2 Оценочные материалы для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине

Вопросы к дифф.зачету

№ п/п	Описание
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	

Образцы задач (заданий) для контрольных (проверочных) работ

Примеры контрольных работ.

Контрольная работа 1. Тест.

1. Междисциплинарность морали как предметной области заключается:

А) В разнообразии социальных практик.

Б) В разнообразии моральных актов и их понимании.

В) В разнообразии моральных теорий.

Г) В разнообразии моральных дилемм.

2. Какая из перечисленных категорий относится к субъективному измерению морального сознания.

А) Совесть.

Б) Добро.

В) Благо.

Г) Идеал.

3. Фраза Марка Аврелия «Делай что должен и будь что будет» по смыслу отражает следующую этическую парадигму

А) Деонтологию.

Б) Аретологию.

В) Консеквенциализм.

Г) Принцип ответственности.

4. Автором принципа ответственности в этике считается

А) И. Кант.

Б) И. Бентам.

В) Аристотель.

Г) Х. Йонас.

5. По мнению представителей какого подхода мораль имеет истоки в поведении животных:

А) Социологического.

Б) Психологического.

В) Онтологического.

Г) Эволюционного.

6. Что является признаком моральной дилеммы:

А) Наличие нескольких вариантов действия.

Б) Связь с ценностями.

В) Нравственные потери или приобретения.

Г) Наличие внешней оценки.

7. Подход основанный на воздании за моральный ущерб называется:

А) Консеквенциализм.

Б) Ретрибутивизм.

В) Гедонизм.

Г) Принцип ответственности.

8. Фраза «Мыслю, следовательно, существую» принадлежит

А) Р. Декарту.

Б) Б. Спинозе.

В) Ф. Бэкону.

Г) Г. В. Лейбницу.

9. Автором мысленного эксперимента «Аргумент Мэри» является

А) А. Тьюринг

Б) Ф. Джексон.

В) Д. Чалмерс.

Г) Х. Патнем.

10. Феномен проявления субъектом своих желаний и намерений, способный к регуляции последующей деятельности и поведения...

- А) Цель.
- Б) Деятельность.
- В) Воля.
- Г) Мотив.

Контрольная работа 2. Эссе.

Темы эссе:

1. Основные возражения к подходу Дж. Серла.
2. Парадигма функционализма Х. Патнема.
3. Моральная машина.
4. Место ИИ в современной правовой практике.
5. Мораль сверхразвитого ИИ.
6. Трансгуманизм.

Примерные вопросы к дифференцированному зачёту.

1. Место морали в жизни человека.
2. Мораль и проблема насилия.
3. Раскройте отношение морали и религии.
4. Основные подходы к понятию моральной ответственности.
5. Основные подходы к эпистемологии морали.
6. Редукционизм и антиредукционизм в понимании сознания.
7. Охарактеризуйте основные подходы к определению сознания (натурализм, дуализм, эпифеноменализм, функционализм).

8. Определение искусственного интеллекта. Тест А. Тьюринга.
9. Мысленный эксперимент "китайская комната" Дж. Серла.
10. Место искусственного интеллекта в современной правовой практике.
11. Искусственный интеллект и этическая экспертиза.
12. Опишите основные перспективы развития ИИ.
13. Перспективы сосуществования ИИ и человека
14. Моральные проблемы расширения разума. Трансгуманизм.

Весь комплект контрольно-измерительных материалов для проверки сформированности компетенции (индикатора компетенции) размещен в закрытой части по адресу, указанному в п. 5.3

6.3 График текущего контроля успеваемости

Неделя	Темы занятий	Вид контроля
4	Тема 3. Основные проблемы философии сознания	
5		
6		Контрольная работа
14	Тема 6. Экзистенциальные вызовы, связанные с внедрением ИИ	
15		
16		Контрольная работа

6.4 Методика текущего контроля

на лекционных занятиях

Текущий контроль включает в себя:

- контроль посещаемости (не менее **80%** занятий), по результатам которого студент получает допуск на дифзачет.

на практических (семинарских) занятиях

Текущий контроль включает в себя

- контроль посещаемости (не менее **80 %** занятий), по результатам которого студент получает допуск на дифзачет,

- выполнение контрольных работ №1, №2

- текущие оценки за работу на семинарах.

Если контрольная работа проводится в виде письменного опроса, то оценка по четырехбалльной шкале выставляется по следующим критериям:

«отлично» - вопросы раскрыты полностью.

«хорошо» - вопросы или отдельный вопрос раскрыт не полностью.

«удовлетворительно» - в ответе на вопросы имеются существенные ошибки;

«неудовлетворительно» - отсутствуют ответы на вопросы или содержание ответа не совпадает с поставленным вопросом.

Работа на семинарах оценивается по четырехбалльной шкале:

«отлично» - глубокое и полное владение материалом.

«хорошо» - изложение материала с несущественными неточностями.

«удовлетворительно» - при ответе допускаются ошибки, владение материалом поверхностное с пробелами, студент не дает ответов на дополнительные вопросы;

«неудовлетворительно» - студент не владеет материалом, допускает грубые ошибки.

В ходе проведения семинарских и практических занятий целесообразно привлечение студентов к как можно более активному участию в дискуссиях, решении задач, обсуждениях и т. д. При этом активность студентов также может учитываться преподавателем, как один из способов текущего контроля на практических занятиях.

самостоятельной работы студентов

Контроль самостоятельной работы студентов осуществляется на лекционных, лабораторных и практических занятиях студентов по методикам, описанным выше.

7 Описание информационных технологий и материально-технической базы

Тип занятий	Тип помещения	Требования к помещению	Требования к программному обеспечению
Лекция	Лекционная аудитория	Количество посадочных мест – в соответствии с контингентом, рабочее место преподавателя, экран, меловая или маркерная доска	
Практические занятия	Аудитория	Количество посадочных мест – в соответствии с контингентом, рабочее место преподавателя, проектор, экран, меловая или маркерная доска.	
Самостоятельная работа	Помещение для самостоятельной работы	Оснащено компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета.	1) Windows XP и выше; 2) Microsoft Office 2007 и выше

8 Адаптация рабочей программы для лиц с ОВЗ

Адаптированная программа разрабатывается при наличии заявления со стороны обучающегося (родителей, законных представителей) и медицинских показаний (рекомендациями психолого-медико-педагогической комиссии). Для инвалидов адаптированная образовательная программа разрабатывается в соответствии с индивидуальной программой реабилитации.

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

№ п/п	Дата	Изменение	Дата и номер протокола заседания УМК	Автор	Начальник ОМОЛА