

Министерство науки и высшего образования Российской Федерации

Томский государственный университет  
систем управления и радиоэлектроники

Г.Н. Нариманова, О.В. Килина

**Этические аспекты искусственного интеллекта и робототехники**

Методические указания к практическим занятиям  
и для выполнения самостоятельной работы обучающихся

Томск  
2024

УДК 621.01  
ББК 334  
Н 28

**Рецензент:**

**Антипин М.Е.**, доцент каф. управления инновациями ТУСУР, канд. физ.-мат. наук

**Нариманова, Гуфана Нурлабековна**

Н 28 Этические аспекты искусственного интеллекта и робототехники: Методические указания к практическим занятиям и для выполнения самостоятельной работы обучающихся по направлению подготовки 15.04.06 «Мехатроника и робототехника»/ Г.Н. Нариманова, О.В. Килина. – Томск: Томск. гос. ун-т систем упр. и радиоэлектронники, 2024. – 16 с.

Методические указания к практическим занятиям и для выполнения самостоятельной работы обучающихся по дисциплине «Этические аспекты искусственного интеллекта и робототехники» разработаны для магистрантов. Методические указания содержат необходимые разъяснения по форме организации практических занятий и самостоятельной работы студентов и ориентированы на достижение результатов образовательной деятельности.

Одобрено на заседании кафедры УИ, протокол № 4 от 21.11.2024.

УДК 621.01  
ББК 334

© Нариманова Г.Н., Килина О.В., 2024  
© Томск. гос. ун-т систем упр. и  
радиоэлектронники, 2024

## Оглавление

1. Введение .....	4
2 Общие требования к проведению практических занятий .....	5
3 Материально-техническое обеспечение практических занятий .....	6
4 Прием результатов выполнения практических заданий .....	7
5 План для практических занятий .....	7
6 Виды самостоятельной работы студентов .....	9
7 Глоссарий дисциплины .....	10
8 Проработка лекционного материала.....	11
9 Примерный перечень вопросов для устного опроса / собеседования .....	12
10 Примерный перечень вопросов для дискуссии .....	13
11 Примерный перечень вопросов для выступления (доклада).....	14
12 Примерный перечень вопросов зачета .....	15
13 Список рекомендуемой литературы .....	16

## 1. Введение

В эпоху невиданного в истории человечества технологического скачка, основанного на развитии искусственного интеллекта (ИИ), интеллектуальной робототехники, вопросы этики – моральных принципов и ценностей, приобретают особую актуальность и значимость. При этом этические соображения в ИИ, робототехнике, автоматизации по мере дальнейшего развития технологий будут только усложняться и обостряться.

Бесспорно, искусственный интеллект значительно превосходит способности человека, который в принципе не в состоянии анализировать и систематизировать гигантские объемы данных, как это делает искусственный интеллект за короткие промежутки времени. Зачастую подлинность такой информации человек проверить объективно не способен, но, между тем, человек предоставляет технологии возможность самостоятельно принимать решения. Возникает вопрос: насколько этичны и безопасны решения и действия искусственного интеллекта?

В связи с этим вполне естественно опасение современного общества относительно того, насколько безопасно столь стремительное распространение ИИ – от использования ИИ в гражданских сферах до его применения в военных разработках, степень безопасности ИИ для человека, в каких областях вообще допустимо его использование, и кто будет нести ответственность за действия искусственного интеллекта. Указанные проблемы в области искусственного интеллекта и робототехники все больше привлекают внимание ученых и практиков во всем мире и побуждают их к активному обсуждению ключевых этических проблем, решение которых, в первую очередь, определяет пути совершенствования регулирования искусственного интеллекта и интеллектуальной робототехники в части морализации технологий, усовершенствования нравственных действий технологий при принятии ими решений. Мир уже однозначно перешел от вопроса о том, следует ли регулировать искусственный интеллект, к неотложному вопросу о том, как это сделать.

Сегодня очень важно работать сообща ученым, инженерам, политикам, специалистам по этике для решения указанных проблем: интегрируя этику на каждом этапе разработки – от этапа проектирования, разработки до внедрения, общество способно создавать ответственные технологии интеллектуальной робототехники, которые приносят пользу человечеству, безусловно, при соблюдении моральных принципов и ценностей.

Правильное управление искусственным интеллектом, осуществление контроля искусственным интеллектом сегодня являются важнейшими задачами, требующими взаимного обучения на основе передового опыта и практики, накопленных в различных юрисдикциях по всему миру.

В контексте образовательной программы по робототехнике и искусственному интеллекту для студентов чрезвычайно важно изучать этические последствия, возникающие при разработке машин, способных принимать самостоятельно решения или выполнять задачи автономно, потому как именно они будут на переднем крае создания технологий, оказывающих значительное влияние на общество.

## 2 Общие требования к проведению практических занятий

Дисциплина «Этические аспекты искусственного интеллекта и робототехники» направлена на изучение вопросов этики применения искусственного интеллекта и робототехники, моральных дилемм современной робототехники, истории этического и правового регулирования ИИ и робототехники в международных документах.

**Ключевые задачи дисциплины** заключаются в формировании у обучающихся представлений:

- о трансформации классических этических проблем в результате развития систем искусственного интеллекта и интеллектуальной робототехники;
- о правовом регулировании создания и использования искусственного интеллекта в Российской Федерации;
- об основных принципах этики и правил поведения (Кодекс этики в сфере искусственного интеллекта);
- о мировом взгляде на искусственный интеллект: различных подходах к этике;
- о перспективах и рисках применения искусственного интеллекта в высокотехнологичных отраслях, образовании, бизнесе.

Практические задания, предусмотренные настоящими указаниями, выполняются студентами во время аудиторных занятий индивидуально или в групповом формате под контролем со стороны преподавателя. Все консультации осуществляются преподавателем. Перед началом занятий студенты должны изучить инструкцию по охране труда. Преподаватель должен убедиться в знании инструкции, задавая студенту вопросы по ее содержанию, после чего сделать соответствующую запись в журнале охраны труда. Во время проведения практических занятий в аудитории студентам запрещается передавать друг другу файлы и другие материалы, являющиеся результатом выполнения заданий. Студент имеет право просить консультации у преподавателя, дополнительных пояснений по теме занятия. Преподаватель, давая консультацию студенту, указывает раздел технической документации или методической литературы, в которой имеется ответ на вопрос студента. Если необходимые сведения в документации и литературе отсутствуют, то преподаватель должен дать устные пояснения или продемонстрировать практические действия, приводящие к требуемому результату, с последующим повторением студентом. Консультации, выдача практических заданий и прием результатов выполнения осуществляется во время аудиторных занятий. При выполнении заданий соблюдается принцип последовательности освоения тем образовательного модуля.

### 3 Материально-техническое обеспечение практических занятий

Для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации используется учебная аудитория с достаточным количеством посадочных мест для учебной группы, оборудованная доской и стандартной учебной мебелью. Имеются мультимедийное оборудование и учебно-наглядные пособия, обеспечивающие тематические иллюстрации по лекционным разделам дисциплины. В случае проведения занятий в режиме онлайн используется необходимое оборудование и программное обеспечение.

Практические занятия должны проводиться в аудитории, оборудованной:

- доской и маркерами;
- проектором и экраном;
- персональными компьютерами, не менее одного на двух студентов группы,

удовлетворяющих требованиям:

- доступом в сеть Internet;
- современным графическим редактором для разработки моделей;
- пакетом офисных приложений для разработки текстов документов.

Размещение и освещенность рабочих мест в учебной аудитории должно удовлетворять действующим требованиям санитарных правил и норм (СанПиН).

#### 4 Прием результатов выполнения практических заданий

Практические занятия являются одной из основных форм образовательного процесса, ориентированной на усвоение обучающимися теоретического материала и выработку практических компетенций. Основной целью практических занятий является комплексный контроль усвоения пройденного материала, хода выполнения студентами самостоятельной работы и рассмотрение наиболее сложных и спорных вопросов в рамках темы учебного занятия. Подготовка к практическим занятиям предполагает самостоятельную работу студентов по изучению материала по конкретной теме.

Результаты выполнения практических заданий демонстрируются преподавателю. Во время приема выполненной работы преподаватель вправе:

- требовать у студента демонстрации выполненного задания в виде файлов, текстов, таблиц, схем, рисунков, в том числе, по возможности и необходимости, в бумажном письменном или распечатанном виде, либо в электронном виде (при размещении результатов выполнения заданий в системе Moodle);

- требовать у студента пояснений, относящихся к способам реализации задания.

Задание считается выполненным и принимается преподавателем только в том случае, если получены все результаты, предусмотренные заданием. Если какие-то результаты, предусмотренные заданием, не получены или неверны, то задание подлежит доработке. Студент должен работать надлежащим образом. Подлежат обязательному исправлению замеченные преподавателем недочеты:

- грамматические ошибки;
- небрежное оформление рисунков, графиков, структур, схем, таблиц;
- неточности в описаниях, структурах, схемах;
- некорректное использование методов, инструментария и т.д.

Результаты выполнения заданий сохраняются студентом в электронном виде (файлы), а также, если возможно и удобно, в бумажном формате, до получения зачета по данной дисциплине. До начала экзаменационной сессии студент обязан сдать результаты выполнения всех практических заданий, предусмотренных настоящим курсом.

#### 5 План для практических занятий

**Модуль 1.** *Искусственный интеллект (ИИ) как фундаментальная наука и технология комплексных технологических решений. Философия и формирование этики ИИ и робототехники. Сознание и моральный статус ИИ*

**Цель занятий:** Освоение понятийного аппарата ИИ, интеллектуальной робототехники, ознакомление с системами ИИ, предпосылками возникновения систем ИИ. Изучение основных направлений исследований в области ИИ и робототехники. Исследование компьютерного понимания естественного языка как важнейшей составляющей моделирования интеллектуальной деятельности человека. Ознакомление с основными приложениями искусственного интеллекта. Формирование у студента понимания о сознании и моральном статусе ИИ, о необходимости разработки ИИ с учетом прав, сознания, самоуважения и свободы, ценности машинного обучения.

**Модуль 2.** *Проблемы правового регулирования создания и использования искусственного интеллекта в Российской Федерации. Перспективы и направления развития этического ИИ в России и мире. Принципы ИИ. Этические аспекты доверия к ИИ*

**Цель занятий:** Формирование представлений у обучающихся о возможных негативных последствиях, возникающих в результате проектирования, разработки и

эксплуатации различными субъектами робототехнических комплексов и систем на базе ИИ. О необходимости регулирования и контроля ИИ на государственном уровне. Этика ИИ в различных междисциплинарных областях: образовании, медицине, машиностроении, строительстве, бизнесе. Ознакомление студентов с основополагающими принципами и ценностями этически-ориентированного ИИ. О профессиональной ответственности проектировщиков и разработчиков систем ИИ, использовании ИИ с целью пропаганды человеческих ценностей и принесения максимальной пользы обществу, о соблюдении законов и нормативно-правовых актов, международных норм поведения. Изучение Кодекса этики в сфере искусственного интеллекта; механизмов его реализации. Ознакомление студентов с понятием доверенного ИИ. Этические аспекты доверия к ИИ. Российские и мировые стандарты, связанные с доверенными системами ИИ.

### **Модуль 3. Технологии ИИ в задачах защиты информации**

**Цель занятий:** Формирование у слушателей представлений об этических аспектах использования ИИ в задачах обеспечения информационной безопасности: проблема защиты персональных данных человека; проблема отчуждения аутентификаторов от личности человека и ее решение с помощью биометрических систем; проблема защиты биометрических персональных данных человека; обезличивание персональных данных. Рассмотрение эффективности использования физиологических и поведенческих признаков человека в задачах идентификации и аутентификации личности. Ознакомление студентов с прикладными вопросами, связанными с развитием ИИ для биометрии. Технологии компьютерного зрения и распознавания образов на базе искусственного интеллекта.

**Модуль 4. Вопросы социальной ответственности, связанные с участием и развитием общества при внедрении систем ИИ. Создание и использование этического и социально приемлемого ИИ в современном мире. Методы оценки рисков, возникающих вследствие внедрения ИИ**

**Цель занятий:** Изучение вопросов по выявлению групп заинтересованных сторон, на которые влияют системы ИИ, и решение их проблем. Минимизация воздействий на окружающую среду при использовании технологий ИИ. Контроль над технологиями ИИ со стороны государства, человека и социума. Рассмотрение опыта применения ИИ в социальном обеспечении. Рынок ИИ в России и мире. Использование ИИ в технологичных отраслях, государственном управлении, бизнесе, образовании и культуре. Влияние использования ИИ на занятость населения за счет повышения автоматизации бизнес-процессов, создания рабочих мест. Изучение возможных рисков использования систем ИИ и пути минимизации их последствий. Международные стандарты, содержащие рекомендации по управлению рисками, с которыми сталкиваются организации при разработке и применении методов и систем ИИ.

## 6 Виды самостоятельной работы студентов

Самостоятельная работа студентов при изучении данной дисциплины предполагает следующие виды работ:

- Подготовку к устному опросу / собеседованию. Предполагает чтение рекомендованной литературы и конспектов лекций, повторение ключевых определений, концепций и примеров, подготовка ответов на возможные вопросы преподавателя.
- Подготовку к дискуссии. Предполагает изучение материалов по теме дискуссии, анализ аргументов, формирование собственной позиции и подготовку контраргументов.
- Подготовку к выступлению (докладу), Подготовка и представление презентаций на различные темы, связанные с курсом, развивает навыки публичных выступлений и структурирования информации.
- Подготовку к зачету. Регулярное выполнение тестов и ответов на контрольные вопросы помогает закрепить полученные знания и выявить пробелы в понимании материала.

Для расширения знаний по дисциплине студенты могут использовать другие формы самостоятельной работы:

- Чтение и анализ научной литературы. Студентам необходимо регулярно знакомиться с новыми исследованиями, статьями и книгами по теме, анализируя ключевые идеи и выводы авторов.
- Анализ реальных ситуаций и проблем, связанных с этикой ИИ и робототехники, требует применения теоретических знаний на практике.
- Самостоятельное исследование отдельных аспектов темы, сбор данных, проведение интервью или опросов, анализ результатов.
- Создание и защита проектов, связанных с применением этических принципов в разработке ИИ-систем или роботов.
- Посещение мероприятий и конференций. Участие в научно-практических конференциях, форумах и выставках, посвященных вопросам ИИ и робототехники, позволяет расширить кругозор и установить профессиональные контакты.
- Участие в дискуссиях и дебатах. Активное участие в обсуждениях и дебатах на семинарах и конференциях способствует развитию коммуникативных навыков и умения аргументированно защищать свою точку зрения.
- Прослушивание вебинаров известных специалистов в области ИИ и робототехники, а также участие в онлайн-дискуссиях после просмотра.
- Написание научных статей. Для магистрантов может быть актуально написание научных статей по результатам своих исследований и публикаций в специализированных журналах.

Эти виды самостоятельной работы помогают студентам лучше усвоить материал, развить аналитическое мышление и подготовиться к профессиональной деятельности в области ИИ и робототехники.

## 7 Глоссарий дисциплины

Автономия робота - способность роботов принимать решения и выполнять задачи без непосредственного вмешательства человека.

Алгоритмическая предвзятость - ситуация, когда алгоритмы ИИ систематически принимают предвзятые решения, основанные на данных, содержащих дискриминацию по признаку пола, расы, возраста и других характеристик.

Бионический протез: Искусственная конечность, имитирующая функции естественной конечности и управляемая сигналами нервной системы.

Виртуальная реальность (VR): Технология создания симулированных сред, которые позволяют пользователям взаимодействовать с виртуальными объектами так, будто они находятся в реальном мире.

Искусственный интеллект (ИИ) - область информатики, которая занимается созданием интеллектуальных машин, способных выполнять задачи, требующие человеческого интеллекта.

Компьютерное зрение: Область исследований, связанная с разработкой методов и алгоритмов, позволяющих компьютерам интерпретировать и понимать визуальную информацию из изображений или видео.

Машинное обучение - подход к искусственному интеллекту, который позволяет компьютерам обучаться на основе данных и улучшать свои результаты без явного программирования.

Обратная связь - процесс получения информации о результатах действий системы ИИ или робота, чтобы корректировать их поведение.

Ответственность за ошибки ИИ - вопрос о том, кто несет ответственность за ущерб, причиненный действиями систем ИИ или роботов.

Персональные данные - информация, относящаяся к идентифицированному или идентифицируемому физическому лицу, которую обрабатывают системы ИИ.

Право на объяснение - право пользователей получать понятное объяснение решений, принятых системами ИИ.

Принципы этического дизайна - набор руководящих принципов, направленных на создание технологий, учитывающих этические соображения.

Роботэтика - раздел прикладной этики, изучающий моральные вопросы, возникающие при разработке и использовании роботов.

Транспарентность ИИ - принцип открытости и доступности информации о работе систем ИИ, включая алгоритмы и данные, используемые для принятия решений.

Угрозы конфиденциальности - риски утечки или неправомерного использования персональных данных, собираемых и обрабатываемых системами ИИ.

## **8 Проработка лекционного материала**

Лекционный материал наряду с рекомендуемой литературой является основой для освоения дисциплины. Составной частью самостоятельной работы по лекционному курсу является непосредственная работа на лекциях – ведение конспектов. Самостоятельная проработка материала прочитанных лекций предполагает изучение конспектов лекций, а также материалов лекций по источникам, приведенным в списке основной и дополнительной учебной литературы. Изучать курс рекомендуется по темам, предварительно ознакомившись с содержанием каждой из них.

## **9 Примерный перечень вопросов для устного опроса / собеседования**

1. Предпосылки возникновения ИИ.
2. Основные приложения искусственного интеллекта.
3. Конфиденциальность данных и личности. Прозрачность в использовании данных. Безопасность больших данных
4. Основные принципы ИИ.
5. Этические аспекты использования ИИ в задачах обеспечения информационной безопасности.
6. Документы нормативно-технического регулирования в сфере ИИ.
7. Прозрачность и объяснимость работы ИИ.
8. Основные понятия инженерной этики.
9. Нано-роботы, роботы, алгоритмы и искусственный интеллект. Этическое и правовое регулирование робототехники в международных документах.
10. Вопросы социальной ответственности, связанные с участием и развитием общества при внедрении систем ИИ

## 10 Примерный перечень вопросов для дискуссии

1. Способность когнитивных роботов к восприятию, к использованию языка, к взаимодействию, к решению проблем, к обучению, к творчеству. Роботы как моральные агенты.
2. Этические принципы для разработки рекомендательных систем в различных областях (медицина, робототехника, государственное регулирование и др.).
3. Роботика и кодексы робототехников. Этические аспекты социального взаимодействия робота и человека.
4. Ответственность частных лиц и организаций при негативных последствиях использования ИИ.
5. Социальная оценка технологий. Принцип «не навреди». Принцип предосторожности.
6. Перспективы развития социальной оценки технологий: эмоциональный дизайн, дизайн ценностей, интернет вещей, оценка воздействия на окружающую среду.
7. Сильный ИИ и слабый ИИ: какова разница в постановке этических проблем. «Может ли машина мыслить? Может ли алгоритм мыслить?».
8. Применение на практике этических принципов разработки и внедрения систем ИИ.
9. Перспективы и направления развития этического ИИ в России и мире.
10. Этические проблемы, вызванные сменой социального ландшафта: исчезновение многих профессий, ускорение темпа жизни, системы социального рейтинга и др.

## 11 Примерный перечень вопросов для выступления (доклада)

1. Внедрение этических принципов ИИ в приоритетные отрасли экономики. Опыт ведущих высокотехнологичных компаний.
2. Особенности этических проблем, связанных с изменением социального ландшафта (по областям).
3. Правовое регулирование искусственного интеллекта, роботов и объектов робототехники.
4. Контроль общества с помощью систем искусственного интеллекта.
5. Акторы в сфере ИИ. Этические стандарты поведения акторов в сфере ИИ.
6. Военное и гражданское использование робототехнических систем. Этика в робототехнике.
7. Создание и использование этичного и социально приемлемого ИИ в современном мире.
8. Эмоциональный ИИ: проблема эмпатии.
9. Этические аспекты безопасности, доверенности и корректности работы автономных интеллектуальных систем.
10. Философия и формирование этики искусственного интеллекта. Машинная мораль: создание или обучение

## 12 Примерный перечень вопросов зачета

1. Этические кодексы мира. Этические кодексы ведущих компаний по разработке систем ИИ
2. Этические проблемы ИИ.
3. Создание и использование этичного и социально приемлемого ИИ в современном мире.
4. Предпосылки возникновения систем искусственного интеллекта.
5. Методы оценки рисков, возникающих из-за внедрения ИИ.
6. Особенности этической экспертизы технологий ИИ.
7. Этические нормы в коммуникации с системами ИИ.
8. Принцип ИИ: справедливость и исключение дискриминации при использовании технологий ИИ для различных социальных групп.
9. Тест Тьюринга для систем искусственного интеллекта.
10. Этические аспекты использования ИИ в задачах обеспечения информационной безопасности: проблема защиты персональных данных

### 13 Список рекомендуемой литературы

1. Гаспарян, Д. Э. Прикладные проблемы внедрения этики искусственного интеллекта в России. Отраслевой анализ и судебная система: монография / Д. Э. Гаспарян, Е. М. Стырин; Нац. исслед. ун-т «Высшая школа экономики». - 2-е изд. - Москва: Изд. дом Высшей школы экономики, 2021. - 112 с. - ISBN 978-5-7598-2242-4. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1910949> (дата обращения: 16.11.2024).
2. Новые законы робототехники. Регуляторный ландшафт. Мировой опыт регулирования робототехники и технологий искусственного интеллекта / В.В. Бакуменко, А. Д. Волынец, А.В. Незнамов [и др.]; под редакцией А. В. Незнамова. — Москва: Infotropic Media, 2018. — 220 с. — ISBN 978-5-9998-0324-5. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/138977> (дата обращения: 16.11.2024).
3. Правовое регулирования искусственного интеллекта, роботов и объектов робототехники как условие формирования экономического лидерства в России: монография / Г. Ф. Ручкина, М. В. Демченко, А. В. Попова [и др.]; под ред. Г.Ф. Ручкиной. - Москва: Прометей, 2021. - 350 с. - ISBN 978-5-00172-197-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1851280> (дата обращения: 16.11.2024).
4. Системы искусственного интеллекта: Учебное пособие / Н. В. Замятин - 2018. 244 с. [Электронный ресурс]: — Режим доступа: <https://edu.tusur.ru/publications/7269> (дата обращения: 17.11.2024).