

Методические рекомендации к курсу «Доверенный  
искусственный интеллект» для направления 09.04.04  
«Программная инженерия»

Никольская Ксения Юрьевна

2021

**Цель дисциплины:** Целью дисциплины является изучение вопросов разработки доверенного искусственного интеллекта.

**Задачи дисциплины:** Задачи дисциплины: получение практических навыков работы с нормативно правовыми документами в области доверенного искусственного интеллекта; получения практических навыков разработки доверенного искусственного интеллекта; научиться оценивать корректность больших данных; научиться оценивать корректность принятия решений доверенным искусственным интеллектом.

**Основная литература по курсу:**

1. Сердюков, Ю. М. Философия виртуальной реальности и искусственного интеллекта: учебное пособие / Ю. М. Сердюков; под редакцией Ю. М. Сердюкова. — Хабаровск: ДВГУПС, 2020. — 169 с. — ISBN 978-5-262-00881-0. — Текст электронный // Лань: электронно-библиотечная система. Режим доступа: для авториз. пользователей. <https://e.lanbook.com/book/179385>

**Дополнительная литература по курсу:**

1. 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. Режим доступа: для авториз. пользователей. <https://e.lanbook.com/book/138977>.

**Объем и виды учебной работы:**

- Семестр: 2.
- Общая трудоёмкость дисциплины: 108 часов.
- Лекции: 16 часов.
- Практические занятия: 16 часов.

**Компетенции:**

1. УК-91 Способен понимать фундаментальные принципы работы современных систем искусственного интеллекта, разрабатывать правила и стандарты взаимодействия человека и искусственного интеллекта и использовать их в социальной и профессиональной деятельности:

*Знает:* правовую базу информационного законодательства, правовые нормы и стандарты в области искусственного интеллекта и смежных областей.

*Умеет:* применять правовые нормы и стандарты в области искусственного интеллекта при создании систем искусственного интеллекта.

2. ПК-4 Способен разрабатывать и модернизировать программное и аппаратное обеспечение технологий и систем искусственного интеллекта с учетом требований информационной безопасности в различных предметных областях:

*Знает:* этические аспекты безопасности, доверенности и корректности работы автономных интеллектуальных и рекомендательных систем.

*Умеет:* ориентироваться в нормативно-правовых документах в области этики искусственного интеллекта.

*Имеет практический опыт:* разработки систем искусственного интеллекта, отвечающим всем современным стандартам в области безопасного искусственного интеллекта.

3. ПК-11 Способен создавать и применять методы объяснимого искусственного интеллекта для создания интерпретируемых интеллектуальных систем:

*Знает:* стандарты и принципы объяснимого искусственного интеллекта.

*Умеет:* применять и разрабатывать стандарты объяснимого искусственного интеллекта, постулирующие принципы прозрачности и объяснимости, чтобы вызывать доверие к своему функционированию и уверенность в выводах системы.

*Имеет практический опыт:* применения стандартов в области объяснимого искусственного интеллекта.

### **Рекомендации к выполнению практических работ:**

Рекомендации к выполнению **Практической работы 1** «Нормативно-правовые документы в области применения доверенного ИИ (Ethics guidelines for trustworthy AI – Европейская Комиссия; Trustworthy AI in health – Саудовская Аравия и др.):

Студенту необходимо детально проанализировать документы в области применения доверенного ИИ, согласно своему варианту.

### **Вопросы для тестирования по усвоению материала практической работы № 1:**

1. Вопросы развития и регулирования искусственного интеллекта в международных организациях.
2. Перечислите основные разработки в области регулирования искусственного интеллекта различными компаниями.
3. Раскройте определение сфер и методов регулирования, категорирование систем.
4. Конкретизируйте применительную практики по регулированию искусственного интеллекта на международном уровне.
5. Вопросы этики и права человека применительно к технологиям искусственного интеллекта.

Рекомендации к выполнению **Практической работы 2** «Методы оценки корректности принятия решений доверенным ИИ»:

Студенту необходимо прописать методы оценки корректности принятия решений доверенным ИИ, согласно своему варианту.

### **Вопросы для тестирования по усвоению материала практической работы № 2:**

1. Объясните взаимосвязь между точность и объяснимостью принятия решений искусственным интеллектом.
2. Дайте определение и объясните суть точности в принятии решений искусственным интеллектом.

3. Дайте определение и объясните суть объяснимости в принятии решений искусственным интеллектом.
4. Какими методами можно предотвратить искажение в системах искусственного интеллекта для распознавания лиц и голоса.
5. Какими методами можно предотвратить искажение в медицинских системах искусственного интеллекта.

Рекомендации к выполнению **Практической работы 3** «Алгоритмы оценки корректности больших данных. Методы защиты больших данных.»:

Необходимо выполнить 2 пункта:

1. Описать пошагово алгоритм оценки корректности больших данных, согласно своему варианту.
2. Описать методы защиты больших данных, согласно своему варианту.

**Вопросы для тестирования по усвоению материала практической работы № 3:**

1. Перечислите принципы защиты распределенных данных.
2. Перечислите принципы защиты нереляционных данных.
3. Перечислите принципы защиты данных в режиме реального времени.
4. Перечислите принципы защиты данных с использованием криптографических средств защиты.
5. Перечислите принципы отслеживания происхождения данных.

Рекомендации к выполнению **Практической работы 4** «Основные принципы разработки автономных систем ИИ в различных областях (беспилотный транспорт, робототехника, медицина и др.):»:

Студенту необходимо расписать основные принципы разработки автономных систем ИИ в различных областях (беспилотный транспорт, робототехника, медицина и др.), согласно своему варианту.

**Вопросы для тестирования по усвоению материала практической работы № 4:**

1. Перечислите основные тенденции развития приложений технологии искусственного интеллекта для интеллектуальных беспилотных автономных систем.
2. Перечислите основные тенденции развития приложений технологии искусственного интеллекта для интеллектуальных беспилотных летательных аппаратов.
3. Перечислите статус гражданских беспилотных летательных аппаратов в России.
4. Перечислите статус гражданских беспилотных летательных аппаратов за рубежом.
5. Перечислите основные тенденции развития приложений технологии искусственного интеллекта для сервисных роботов.

Рекомендации к выполнению **Практической работы 5**  
«Стандартизация разработки доверенного ИИ (ISO/IEC AWI 5339; ISO/IEC DTS 4213 и др.)»:

Студенту необходимо описать стандарт разработки доверенного ИИ, согласно своему варианту.

**Вопросы для тестирования по усвоению материала практической работы № 5:**

1. Дайте определение стандартизации и ее назначение в области разработки систем доверенного искусственного интеллекта.
2. Перечислите основополагающие стандарты в области искусственного интеллекта.
3. Перечислите основополагающие стандарты в области больших данных.
4. Перечислите основополагающие стандарты в области надежности искусственного интеллекта.
5. Перечислите основополагающие стандарты в области вычислительных подходов и вычислительных характеристик систем искусственного интеллекта.

Рекомендации к выполнению **Практической работы 6** «Основные этапы проектирования, разработки и внедрения доверенного ИИ на реальных примерах»:

Необходимо прописать основные этапы проектирования, разработки и внедрения доверенного ИИ на реальных примерах, согласно своему варианту.

**Вопросы для тестирования по усвоению материала практической работы № 6:**

1. Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области медицины.
2. Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области защиты сетевых данных.
3. Перечислите основные этапы проектирования систем доверенного искусственного интеллекта, работающих с биометрическими данными.
4. Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области образования.
5. Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области рекомендательных систем.

Рекомендации к выполнению **Практической работы 7** «Законодательное регулирование и риски внедрения и использования доверенного ИИ»:

Необходимо описать как законодательно регулируется и какие риски внедрения и использования доверенного ИИ, согласно своему варианту.

**Вопросы для тестирования по усвоению материала практической работы № 7:**

1. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в России в контексте использования систем доверенного искусственного интеллекта.

2. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в Китае в контексте использования систем доверенного искусственного интеллекта.
3. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в Австралии в контексте использования систем доверенного искусственного интеллекта.
4. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в Евросоюзе в контексте использования систем доверенного искусственного интеллекта.
5. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в США в контексте использования систем доверенного искусственного интеллекта.



**Рекомендации по подготовке к тестированиям по усвоению материалов лекций:**

**Рекомендации по подготовке к тестированию по усвоению материалов лекции № 1 «Введение в доверенный ИИ»:**

**Вопросы:**

1. Дайте определение доверенного искусственного интеллекта.
2. Перечислите существующие концепции доверия.
3. Опишите универсальную модель для создания доверенного искусственного интеллекта.
4. Перечислите и опишите основные принципы надежного искусственного интеллекта.
5. Опишите ключевые аспекты, которые должны присутствовать в системах доверенного искусственного интеллекта.

**Рекомендации по подготовке к тестированию по усвоению материалов лекции № 2 «Нормативно-правовые документы в области применения доверенного ИИ (Ethics guidelines for trustworthy AI – Европейская Комиссия; Trustworthy AI in health – Саудовская Аравия и др.)»:**

**Вопросы:**

1. Перечислите основные нормативно-правовые документы, существующие в настоящее время.
2. Перечислите основные принципы и дайте им характеристику, которые должны присутствовать в доверенных системах искусственного интеллекта в здравоохранении.
3. Перечислите основные принципы нормативно-правовых актов в области разработки и применения доверенного искусственного интеллекта в США.
4. Перечислите основные принципы нормативно-правовых актов в области разработки и применения доверенного искусственного интеллекта в России.

5. Перечислите основные принципы нормативно-правовых актов в области разработки и применения доверенного искусственного интеллекта в Австралии.

Рекомендации по подготовке к тестированию **по усвоению материалов лекции № 3 «Методы оценки корректности принятия решений доверенным ИИ»:**

**Вопросы:**

1. Перечислите основные требования к процессам проектирования и разработки систем доверенного искусственного интеллекта.
2. Объясните принцип прозрачности при работе систем доверенного искусственного интеллекта.
3. Объясните принцип точности при работе систем доверенного искусственного интеллекта.
4. Объясните принцип воспроизводимости при работе систем доверенного искусственного интеллекта.
5. Объясните принцип справедливости при работе систем доверенного искусственного интеллекта.

Рекомендации по подготовке к тестированию **по усвоению материалов лекции № 4 «Алгоритмы оценки корректности больших данных. Методы защиты больших данных.»:**

**Вопросы:**

1. Перечислите оценку достоверности онлайн данных.
2. Перечислите принципы создания обучающих наборов данных, состоящих из изображений, для систем доверенного искусственного интеллекта.
3. Перечислите принципы создания обучающих наборов данных, состоящих из статистической информации, для систем доверенного искусственного интеллекта.

4. Перечислите принципы создания обучающих наборов данных, состоящих из аудиоданных, для систем доверенного искусственного интеллекта.

5. Перечислите методы обеспечения безопасности больших данных.

Рекомендации по подготовке к тестированию **по усвоению материалов лекции № 5 «Основные принципы разработки автономных систем ИИ в различных областях (беспилотный транспорт, робототехника, медицина и др.)»:**

**Вопросы:**

1. Дайте определение автономной системы искусственного интеллекта.
2. Перечислите основные принципы разработки беспилотных летательных аппаратов.
3. Перечислите основные принципы разработки роботизированных систем.
4. Перечислите основные принципы разработки автономных систем искусственного интеллекта в медицине.
5. Перечислите основные принципы разработки автономных систем искусственного интеллекта в образовании.

Рекомендации по подготовке к тестированию **по усвоению материалов лекции № 6 «Стандартизация разработки доверенного ИИ (ISO/IEC AWI 5339; ISO/IEC DTS 4213 и др.)»:**

**Вопросы:**

1. Перечислите стандарты в области разработки доверенного искусственного интеллекта.
2. Что представляет собой процесс стандартизации.
3. Перечислите основные принципы стандартизации при проведении обучения систем доверенного искусственного интеллекта.

4. Перечислите основные принципы стандартизации при создании обучающих наборов данных для систем доверенного искусственного интеллекта.
5. Перечислите основные принципы при проведении аудита систем доверенного искусственного интеллекта.

**Рекомендации по подготовке к тестированию по усвоению материалов лекции № 7 «Основные этапы проектирования, разработки и внедрения доверенного ИИ на реальных примерах»:**

**Вопросы:**

1. Перечислите основные этапы проектирования систем с доверенным искусственным интеллектом.
2. Перечислите основные этапы разработки систем с доверенным искусственным интеллектом.
3. Перечислите основные этапы внедрения систем с доверенным искусственным интеллектом.
4. Перечислите основные этапы аудита систем с доверенным искусственным интеллектом.
5. Дайте определение совместного дизайна в контексте проектирования систем с доверенным искусственным интеллектом.

**Рекомендации по подготовке к тестированию по усвоению материалов лекции № 8 «Законодательное регулирование и риски внедрения и использования доверенного ИИ»:**

**Вопросы:**

1. Перечислите основные принципы, которые необходимы для регулирования технологии распознавания лиц.
2. Перечислите основные принципы, которые необходимы для регулирования автономного вождения.
3. Перечислите основные принципы, которые необходимы для регулирования рекомендательных систем.

4. Перечислите основные принципы, которые необходимы для регулирования приложений, обеспечивающих сетевую безопасность.
5. Перечислите основные принципы, которые необходимы для регулирования систем доверенного искусственного интеллекта, которые работают с персональными данными.

### **Вопросы к зачету:**

1. Перечислите основные принципы, которые необходимы для регулирования технологии распознавания лиц.
2. Перечислите основные принципы, которые необходимы для регулирования автономного вождения.
3. Перечислите основные принципы, которые необходимы для регулирования рекомендательных систем.
4. Перечислите основные принципы, которые необходимы для регулирования приложений, обеспечивающих сетевую безопасность.
5. Перечислите основные принципы, которые необходимы для регулирования систем доверенного искусственного интеллекта, которые работают с персональными данными.
6. Перечислите основные этапы проектирования систем с доверенным искусственным интеллектом.
7. Перечислите основные этапы разработки систем с доверенным искусственным интеллектом.
8. Перечислите основные этапы внедрения систем с доверенным искусственным интеллектом.
9. Перечислите основные этапы аудита систем с доверенным искусственным интеллектом.
10. Дайте определение совместного дизайна в контексте проектирования систем с доверенным искусственным интеллектом.
11. Перечислите стандарты в области разработки доверенного искусственного интеллекта.
12. Что представляет собой процесс стандартизации.
13. Перечислите основные принципы стандартизации при проведении обучения систем доверенного искусственного интеллекта.
14. Перечислите основные принципы стандартизации при создании обучающих наборов данных для систем доверенного искусственного интеллекта.

15. Перечислите основные принципы при проведении аудита систем доверенного искусственного интеллекта.
16. Дайте определение автономной системы искусственного интеллекта.
17. Перечислите основные принципы разработки беспилотных летательных аппаратов.
18. Перечислите основные принципы разработки роботизированных систем.
19. Перечислите основные принципы разработки автономных систем искусственного интеллекта в медицине.
20. Перечислите основные принципы разработки автономных систем искусственного интеллекта в образовании.
21. Перечислите оценку достоверности онлайн данных.
22. Перечислите принципы создания обучающих наборов данных, состоящих из изображений, для систем доверенного искусственного интеллекта.
23. Перечислите принципы создания обучающих наборов данных, состоящих из статистической информации, для систем доверенного искусственного интеллекта.
24. Перечислите принципы создания обучающих наборов данных, состоящих из аудиоданных, для систем доверенного искусственного интеллекта.
25. Перечислите методы обеспечения безопасности больших данных.
26. Перечислите основные требования к процессам проектирования и разработки систем доверенного искусственного интеллекта.
27. Объясните принцип прозрачности при работе систем доверенного искусственного интеллекта.
28. Объясните принцип точности при работе систем доверенного искусственного интеллекта.
29. Объясните принцип воспроизводимости при работе систем доверенного искусственного интеллекта.

30. Объясните принцип справедливости при работе систем доверенного искусственного интеллекта.
31. Перечислите основные нормативно-правовые документы, существующие в настоящее время.
32. Перечислите основные принципы и дайте им характеристику, которые должны присутствовать в доверенных системах искусственного интеллекта в здравоохранении.
33. Перечислите основные принципы нормативно-правовых актов в области разработки и применения доверенного искусственного интеллекта в США.
34. Перечислите основные принципы нормативно-правовых актов в области разработки и применения доверенного искусственного интеллекта в России.
35. Перечислите основные принципы нормативно-правовых актов в области разработки и применения доверенного искусственного интеллекта в Австралии.
36. Дайте определение доверенного искусственного интеллекта.
37. Перечислите существующие концепции доверия.
38. Опишите универсальную модель для создания доверенного искусственного интеллекта.
39. Перечислите и опишите основные принципы надежного искусственного интеллекта.
40. Опишите ключевые аспекты, которые должны присутствовать в системах доверенного искусственного интеллекта.
41. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в России в контексте использования систем доверенного искусственного интеллекта.
42. Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в Китае в контексте использования систем доверенного искусственного интеллекта.



- 43.Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в Австралии в контексте использования систем доверенного искусственного интеллекта.
- 44.Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в Евросоюзе в контексте использования систем доверенного искусственного интеллекта.
- 45.Перечислите основные нормативно-правовые акты, которые регулируют работу с персональными данными в США в контексте использования систем доверенного искусственного интеллекта.
- 46.Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области медицины.
- 47.Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области защиты сетевых данных.
- 48.Перечислите основные этапы проектирования систем доверенного искусственного интеллекта, работающих с биометрическими данными.
- 49.Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области образования.
- 50.Перечислите основные этапы проектирования систем доверенного искусственного интеллекта в области рекомендательных систем.
- 51.Дайте определение стандартизации и ее назначение в области разработки систем доверенного искусственного интеллекта.
- 52.Перечислите основополагающие стандарты в области искусственного интеллекта.
- 53.Перечислите основополагающие стандарты в области больших данных.
- 54.Перечислите основополагающие стандарты в области надежности искусственного интеллекта.
- 55.Перечислите основополагающие стандарты в области вычислительных подходов и вычислительных характеристик систем искусственного интеллекта.

- 56.Перечислите основные тенденции развития приложений технологии искусственного интеллекта для интеллектуальных беспилотных автономных систем.
- 57.Перечислите основные тенденции развития приложений технологии искусственного интеллекта для интеллектуальных беспилотных летательных аппаратов.
- 58.Перечислите статус гражданских беспилотных летательных аппаратов в России.
- 59.Перечислите статус гражданских беспилотных летательных аппаратов за рубежом.
- 60.Перечислите основные тенденции развития приложений технологии искусственного интеллекта для сервисных роботов.
- 61.Перечислите принципы защиты распределенных данных.
- 62.Перечислите принципы защиты нереляционных данных.
- 63.Перечислите принципы защиты данных в режиме реального времени.
- 64.Перечислите принципы защиты данных с использованием криптографических средств защиты.
- 65.Перечислите принципы отслеживания происхождения данных.
- 66.Объясните взаимосвязь между точность и объяснимостью принятия решений искусственным интеллектом.
- 67.Дайте определение и объясните суть точности в принятии решений искусственным интеллектом.
- 68.Дайте определение и объясните суть объяснимости в принятии решений искусственным интеллектом.
- 69.Какими методами можно предотвратить искажение в системах искусственного интеллекта для распознавания лиц и голоса.
- 70.Какими методами можно предотвратить искажение в медицинских системах искусственного интеллекта.
- 71.Вопросы развития и регулирования искусственного интеллекта в международных организациях.

- 72.Перечислите основные разработки в области регулирования искусственного интеллекта различными компаниями.
- 73.Раскройте определение сфер и методов регулирования, категорирование систем.
- 74.Конкретизируйте применительную практики по регулированию искусственного интеллекта на международном уровне.
- 75.Вопросы этики и права человека применительно к технологиям искусственного интеллекта.