Конспект лекции

С.С. Жумажанова, П.С. Ложников

Этика и правовые проблемы искусственного интеллекта

**Национальный кодекс этики в сфере искусственного интеллекта**

**НАЦИОНАЛЬНЫЙ КОДЕКС ЭТИКИ В СФЕРЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

**1 Необходимость создания Кодекса этики в сфере ИИ**

Современные ИТ-технологии относятся к слабому искусственному интеллекту. Однако, скорость их развития позволяет судить о том, что процесс перехода к сильному искусственному интеллекту возможен в ближайшем будущем. Учитывая расширение сфер внедрения технологий на базе ИИ рамки его применения будут распространяться на те сферы жизнедеятельности человека, для которых необходимо проводить этический анализ, например, в медицине. Поэтому разработка определенных этических норм по созданию и использованию ИИ является крайней необходимостью.

**2 Содержание Кодекса**

Раздел 1. Принципы этики и правила поведения.

1. Главный приоритет развития технологий ИИ в защите интересов и прав людей и отдельного человека.

2. Необходимо осознавать ответственность при создании и использовании ИИ.

3. Ответственность за последствия применения СИИ всегда несет человек.

4. Технологии ИИ нужно применять по назначению и внедрять там, где это принесет пользу людям.

5. Интересы развития технологий ИИ выше интересов конкуренции.

6. Важна максимальная прозрачность и правдивость в информировании об уровне развития технологий ИИ, их возможностях и рисках.

Раздел 2. Применение кодекса.

1. Основы действия Кодекса.

2. Механизм присоединения и реализации кодекса.

**3 Акторы в сфере ИИ**

Акторами в сфере искусственного интеллекта понимается круг лиц, в том числе иностранных, принимающих участие в жизненном цикле СИИ, при его реализации на территории Российской Федерации или в отношении лиц, находящихся на территории Российской Федерации, включая предоставление товаров и оказание услуг.

(Акторы ИИ) относятся, в том числе:

* разработчики, создающие, обучающие, тестирующие модели/системы ИИ и разрабатывающие, реализующие такие модели/системы, программные и/или аппаратные комплексы и принимающие на себя ответственность в отношении их конструкции;
* заказчики (лицо или организация), получающие продукт или услугу;
* поставщики данных и лица, осуществляющие формирование наборов данных для применения их в СИИ;
* эксперты, осуществляющие измерение и/или оценку параметров разработанных моделей/систем;
* изготовители, осуществляющие производство СИИ;
* эксплуатанты СИИ, на законном основании владеющие соответствующими системами, использующие их по назначению и непосредственно реализующие решение прикладных задач с использованием СИИ;
* операторы (лицо или организация), осуществляющие работу СИИ;
* лица, принимающие участие в регуляторном воздействии на сферу ИИ, в том числе разработчики нормативно-технических документов, руководств, различных регуляторных положений, требований и стандартов в области ИИ;
* иные лица, действия которых потенциально могут повлиять на результаты действий СИИ или лиц, принимающих решения с использованием СИИ.

**4 Общие этические принципы/ценности поведения Акторов ИИ**

* Человеко-ориентированный и гуманистический подход.
* Риск-ориентированный подход.
* Принцип предосторожности.
* Ответственное отношение.

**5 Этические стандарты поведения Акторов в сфере ИИ**

* Действовать ответственно и добросовестно
* Соблюдать интересы общества и граждан
* Проявлять профессионализм и компетентность
* Содействовать доверию общества к технологиям ИИ в интересах устойчивого развития

**6 Механизмы реализации Кодекса этики в сфере ИИ**

* Оценка рисков.
* Уполномоченный по этике и комиссии.
* Свод наилучших практик и реестр.
* Повышение осведомлённости об этике применения ИИ.
* Добровольная сертификация.
* Пересмотр положений Кодекса.

**7 Документы нормативно-технического регулирования в сфере ИИ**

* + - ГОСТ Р ИСО 26000-2012, «Руководство по социальной ответственности», идентичен международному стандарту ISO 26000:2010 "Guidance on social responsibility" (раскрывает основные проблемы и принципы социальной ответственности, содержит определения терминов «подотчетность», «должная предусмотрительность», «этичное поведение», «прозрачность» и др.);
    - ГОСТ Р 54089-2010. «Интегрированная логистическая поддержка. Электронное дело изделия. Основные положения» (содержит определения «разработчики», «изготовители», «эксплуатанты»);
    - ГОСТ Р ИСО 9001-2015 Системы менеджмента качества. Требования. Переиздание, февраль 2020 г., идентичен международному стандарту ISO 9001:2015 "Quality management systems - Requirements", IDT (приведены принципы менеджмента качества, раскрыта концепция «риск-ориентированного мышления»);
    - ГОСТ Р ИСО 31000—2019 «Менеджмент риска: принципы и руководство», идентичен международному стандарту ISO 31000:2018 "Risk management - Guidelines", IDT (содержит определения «риск», «менеджмент риска», «управление риском»);
    - ГОСТ Р ИСО 9000-2015 Системы менеджмента качества. Основные положения и словарь (Издание с Поправкой), 6 издание, октябрь 2019 г. (раскрыты понятия «Качество», «Система менеджмента качества», «Принятие решений, основанное на свидетельствах»);
    - ГОСТ Р ИСО/МЭК 12207-2010 Информационная технология. Системная и программная инженерия. Процессы жизненного цикла программных средств (определения «жизненный цикл», «аудит», «заказчик», «разработчик», «оценивание», «исполнитель», «система»);
    - ГОСТ Р 53625-2009 (ИСО/МЭК 19796-1:2005) Информационная технология (ИТ). Обучение, образование и подготовка. Менеджмент качества, обеспечение качества и метрики. Часть 1. Общий подход. Переиздание, октябрь 2018 г. (раскрыт общий подход к качеству на основе основополагающих принципов и требований);
    - ГОСТ Р 53723-2009 Информационные технологии (ИТ). Руководство по применению ГОСТ Р 53625-2009, Переиздание, октябрь 2018 г. (раскрывает подходы к обеспечению качества электронного обучения, приведены «эталонные критерии качества (ЭКК)»);
    - ГОСТ Р ИСО/МЭК ТО 19791-2008. Информационная технология. Методы и средства обеспечения безопасности. Оценка безопасности автоматизированных систем (определения «менеджмент рисков», «верификация»);
    - ГОСТ Р 52653-2006 Информационно-коммуникационные технологии в образовании. Термины и определения. Переиздание, октябрь 2018 г. (приведены определения терминов «информация», «информационная технология», «информационная система»);
    - ГОСТ Р ИСО/МЭК 9126-93 Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению (определение «программно-аппаратные средства», «качество», «уровень качества функционирования»);
    - ГОСТ 20886-85 Группа Э00 Организация данных в системах обработки данных. Термины и определения (определения: «набор данных», «защита данных»).
* Указ Президента РФ № 490 от 10.10.2019 г. «О развитии искусственного интеллекта в Российской Федерации».
* Распоряжение Правительства №2129-р от 19.08.2020 об утверждении «Концепции регулирования искусственного интеллекта и робототехники до 2024 года».
* ГОСТ Р 59276-2020 «Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения».
* Федеральный закон "О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации - городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона "О персональных данных" от 24.04.2020 № 123-ФЗ.

**Список источников**

1. Кодекс этики в сфере искусственного интеллекта [electronic resource]. URL: <https://a-ai.ru/wp-content/uploads/2021/10/кодекс_этики_в_сфере_ии_финальный.pdf>.