

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Локтионова Оксана Геннадьевна

Должность: проректор по учебной работе

Дата подписания: 26.07.2022 10:13:58

Уникальный программный ключ:

0b817ca911e6668abb13a5b426d39e5f1c1eabb73e943df4a4851fda56d089

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
ВОЛГОГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ

На правах рукописи

**Драгунов Станислав Евгеньевич
Матохина Анна Владимировна**

**Проработка этических и правовых вопросов искусственного
интеллекта в выбранной области.**

*Методические указания для лабораторных и практических работ по
дисциплине
«Системы искусственного интеллекта»*

Волгоград 2021

Задание на практическую работу: в соответствии с выбранным заданием по проектированию мультиагентной системы в одной из сфер, проработать этические и правовые вопросы.

В соответствии со стандартами IEEE P7000:

1. Выявить и проанализировать потенциальные этические проблемы в системе или программном обеспечении, с целью избежания негативных и непредвиденных последствий:
 - a. определить критерии для оценки прозрачности для заинтересованных сторон
 - b. определить критерии для оценки и методы измерения средств контроля конфиденциальности в системе.
 - c. рассмотреть возможные варианты алгоритмических предубеждений, для избежания ситуаций отрицательной предвзятости (наборы или информация, заведомо несовместимые с законодательством или со случаями предвзятого отношения снижающим благосостояние заинтересованных сторон или пользователей), а также явные или скрытые внушения или манипуляции, призванные повлиять на поведение или эмоции пользователя.
 - d. определить критерии оценки воздействия А / IS на благополучие человека с учетом текущих индексов благополучия для улучшения, управление рисками, оценки производительности и выявление непреднамеренных действий пользователей.
 - e. определить критерии оценки эффективности эмпатической технологии, включая системы, способные определять и эмитировать эмоции.
2. Дать описание требований к системе с точки зрения этических и правовых аспектов с учетом целевой аудитории, назначения А / IS и требований стандартов.
3. Проработать аспекты стандартизации функционала и методов верификации в соответствии со стандартами.
4. Разработать инструкции проверки автономных и интеллектуальных систем в соответствии с передовыми практиками.

Общее описание стандартов

Серия разрабатываемых проектов стандартов IEEE P7000™ представляет собой уникальное дополнение к коллекции из более чем 1900 глобальных стандартов и проектов IEEE. В то время как более традиционные стандарты ориентированы на совместимость технологий, функциональность, безопасность и упрощение процедур торговли, серия IEEE P7000 решает конкретные вопросы на стыке технологических и этических соображений. Как и его аналоги по техническим стандартам, серия IEEE P7000 расширяет границы возможностей инноваций и приносит пользу обществу.

Программа сертификации по этике для автономных и интеллектуальных систем (ЕСРАIS) ставит своей целью создание спецификаций для процессов сертификации и маркировки, которые повышают прозрачность, подотчетность и снижение алгоритмической предвзятости в автономных и интеллектуальных системах. ЕСРАIS намеревается предложить процесс и определить ряд знаков, по которым организации могут запрашивать сертификаты для своих процессов, связанных с продуктами, системами и услугами А / IS, которые они предоставляют.

Целью Программы сертификации по этике для автономных и интеллектуальных систем (ЕСРАIS) является создание спецификаций для процессов сертификации и маркировки, которые повышают прозрачность, подотчетность и снижение алгоритмической предвзятости в автономных и интеллектуальных системах (AIS).

Значение этого процесса сертификации для рынка и общества в целом нельзя недооценивать. Распространение систем в виде умных домов, роботов-компаньонов, автономных транспортных средств или любых бесчисленных продуктов и услуг, которые уже существуют сегодня, отчаянно нуждаются в простой и визуальной коммуникации с потребителями и гражданами, независимо от того, считаются ли они «безопасными» или «заслуживающими доверия». всемирно признанная группа экспертов, предоставляющая общедоступную и прозрачную серию знаков.

Мировой рынок автономных и интеллектуальных систем (А / IS) стремительно растет. По оценкам компаний, занимающихся исследованиями рынка, выручка от решений А / IS вырастет с 19 миллиардов долларов в 2018 году до более чем 52 миллиардов долларов в 2021 году. Уже сегодня потенциальная ценность AI / S в отраслях составляет от 3,5 до 5,8 трлн долларов в год. А / IS широко используется в жизненно важных областях, таких как здравоохранение, оборона или мобильность, а также в большинстве критически важных бизнес-приложений во всех отраслях промышленности будут полагаться на решения А / IS. Доступ к А / IS определяет перспективы этих предприятий сейчас и в будущем.

Хотя А / IS набирает обороты во всех сферах жизни, все большее число экспертов по А / IS и общество в целом обеспокоено этикой и подотчетностью А / IS. Доверие станет важным инструментом для реализации ценности бизнеса А / IS и выхода на рынок, особенно в жизни критические приложения. Ответственность за операции А / IS оценивается клиентами, партнеров, органов власти и СМИ, и становится неотъемлемой частью корпоративной обязанности.

Интеллектуальные приложения будут все больше работать и принимать решения и транзакции без вмешательства человека. Люди несут ответственность за свои решения - каждый А / IS приложение тоже должно быть. Технологическая экосистема требует решений, которые безопасное честное и прозрачное развитие А / IS. Чтобы удовлетворить эти требования, будущее А / IS нуждается в сертифицируемых процессах, проверенных надежной организацией, чтобы легко установить опознаваемые знаки, обеспечивающие их надежность и безопасность для широкой публики.

В нынешней атмосфере развития А / IS существует ряд компаний, организации и академические учреждения, предоставляющие оценки воздействия или другие аналогичные инструменты, помогающие определить, когда продукт, услуга или система подотчетны, заслуживают доверия или выгодный.

Однако на данный момент нет доверенного SDO (организации по разработке стандартов) предложение любой формы сертификации Инициативы, которую организация, использующая влияние А / IS оценка их продуктов, услуг или систем может подойти для получения значка или чтобы продемонстрировать своим клиентам, заинтересованным сторонам и обществу в целом, что они были официально и публично признаны подотчетными, заслуживающими доверия или полезными авторитетный экспертный совет коллег.

Цель программы сертификации по этике для автономных и интеллектуальных систем (ЕСРАИС) ЕСРАИС предоставляет первую в мире (на основе нашего исследования) спецификацию и своего рода тело, позволяющее использовать значок или знак для продуктов, услуг и систем А / IS. В частности, ЕСРАИС позволит проводить оценку на основе процессов и результатов продукты / услуги и системы организации с использованием подхода, основанного на оценке риска.

Фокусные отрасли:

- Государственные учреждения
- Города
- Общественная безопасность
- Здоровье
- Мобильность

- образование
- Телекоммуникации
- Финансовые услуги
- Розничная торговля
- Потенциальные поставщики услуг по сертификации ИИ
- Производство
- Правосудие

Критерии сертификации, ориентированные на этику искусственного интеллекта, предназначены для того, чтобы позволить городам, государственным и частным организациям в различных вертикальных отраслях, таких как здравоохранение и медицинское оборудование, финансовые услуги, автомобилестроение, производство и услуги для пожилых людей, повысить доверие и выгоды использования, услуг и систем А / IS, которые они разрабатывают или используют. Разработка критериев этической прозрачности, подотчетности и конфиденциальности для приложений отслеживания контактов / технологий отслеживания контактов (СТА / СТТ) полностью соответствует цели IEEE по продвижению технологий на благо человечества. ЕСРАИС работает в связке с программой ICAP и предлагает дополнительный механизм для ICAP, чтобы применить свой опыт на соответствующем и необходимом рынке.

ЕСРАИС работает в рамках управляемого отраслью итеративного процесса разработки сертификатов / знаков для А / IS. Цель ЕСРАИС - обеспечить возможность работы в циклах разработки и отраслевой проверки, и были получены следующие результаты:

- Критерии и процесс сертификации прозрачности в А / IS
- Критерии и процесс сертификации подотчетности в А / IS
- Критерии и процесс сертификации алгоритмической предвзятости в А / IS
- Критерии и процесс сертификации конфиденциальности в А / IS

Индивидуальные наборы вышеуказанных критериев, адаптированные к требованиям и приоритетам конкретных отраслей промышленности и областей применения.

IEEE P2089 - Проект стандарта IEEE для концепции цифровых услуг, соответствующих возрасту - на основе принципов 5 прав для детей.

Стандарт устанавливает набор процессов, с помощью которых организации стремятся сделать свои услуги соответствующими возрасту. Стандарт призван поддержать растущее желание организаций разрабатывать цифровые продукты и

услуги с учетом интересов детей и отражает их существующие права в соответствии с Конвенцией Организации Объединенных Наций о правах ребенка (Конвенция). Хотя в разных юрисдикциях могут действовать разные законы и правила, этот стандарт предлагает передовой опыт разработки цифровых услуг, прямо или косвенно влияющих на детей. В нем описываются процессы на протяжении жизненного цикла разработки, доставки и распространения, которые помогут организациям задавать правильные актуальные вопросы о своих услугах. выявлять риски и возможности, с помощью которых можно сделать свои услуги соответствующими возрасту, и принимать меры по снижению рисков и внедрять полезные системы, которые поддерживают более активное участие в соответствующих возрастных категориях. Каждый третий пользователь в сети моложе 18 лет, что означает, что этот стандарт имеет широкое применение. Возраст совершеннолетия определяется статьей 1 Конвенции о правах ребенка.

IEEE P2247™ - Стандарт для адаптивных учебных систем

Этот стандарт определяет и классифицирует компоненты и функциональные возможности адаптивных учебных систем (AIS). Он определяет параметры, используемые для описания AIS, и устанавливает требования и руководство по использованию и измерению этих параметров. Определяет взаимодействие и обмен между компонентами адаптивных систем обучения (AIS). Он определяет данные и структуры данных, используемые в этих взаимодействиях и обменах, и параметры, используемые для их описания и измерения, а также устанавливает требования и руководство по использованию и измерению данных, структур данных и параметров. Предлагает рекомендуемую практику и классифицирует методы оценки адаптивных учебных систем (AIS) и устанавливает руководство по использованию этих методов. Эта передовая практика включает и продвигает принципы этически согласованного дизайна для использования искусственного интеллекта (AI) в AIS.

IEEE P2660.1 Рекомендуемые методы работы с промышленными агентами: интеграция программных агентов и функций автоматизации низкого уровня

В этой рекомендуемой практике описывается интеграция и развертывание технологии многоагентных систем (MAS) в промышленных средах для использования при создании уровня интеллектуального принятия решений. поверх устаревших промышленных платформ управления. Также возможна интеграция

программных агентов с низкоуровневыми системами управления в реальном времени, в основном на основе программируемых логических контроллеров (ПЛК), на которых выполняются программы управления IEC 61131-3™ (образующие таким образом новый компонент, известный как промышленные агенты), идентифицированы. Кроме того, описывается интеграция программных агентов с управляющими приложениями, основанными на стандарте IEC 61499™ или выполняемыми на встроенных контроллерах.

Эта рекомендуемая практика поддерживает и помогает инженерам использовать передовой опыт разработки промышленных агентов для решения конкретных задач автоматизации управления и заданных областей применения. Поэтому предоставляются соответствующие правила, рекомендации и шаблоны проектирования.

IEEE P2671 Стандарт общих требований к онлайн-обнаружению на основе машинного зрения в интеллектуальном производстве

Этот стандарт определяет общие требования к онлайн-обнаружению на основе машинного зрения, включая требования к формату данных, процессам передачи данных, определению сценариев приложений и метрикам производительности. для оценки эффекта от развертывания онлайн-обнаружения.

IEEE P2672 Руководство по общим требованиям массовой настройки

Это руководство содержит определения, терминологию, рабочие процедуры, системные архитектуры, ключевые технологические требования, требования к данным и приложениям, относящиеся к массовой настройке, ориентированной на пользователя. Это руководство содержит справочную информацию, которая будет использоваться производственными предприятиями для разработки и внедрения бизнес-моделей массовой настройки.

IEEE P2801™ - Рекомендуемая практика управления качеством наборов данных для медицинского искусственного интеллекта

В этой рекомендуемой практике определяются передовые методы создания системы управления качеством для наборов данных, используемых для медицинских устройств с искусственным интеллектом. Он охватывает полный цикл управления

наборами данных, включая такие элементы, как, помимо прочего, сбор, передачу, использование, хранение, обслуживание и обновление данных.

Эта рекомендуемая практика рекомендует список критических факторов, которые влияют на качество наборов данных, таких как, помимо прочего, источники данных, качество данных, аннотации, защита конфиденциальности, квалификация / обучение / оценка персонала, инструменты, оборудование, среда, управление процессами и документация. .

IEEE P2802 Стандарт оценки производительности и безопасности медицинского устройства на основе искусственного интеллекта: Терминология.

Этот стандарт устанавливает терминологию, используемую в медицинских устройствах с искусственным интеллектом, включая определения фундаментальных концепций и методологий, которые описывают безопасность, эффективность, риски и управление качеством искусственного интеллекта. разведывательные медицинские приборы.

Он предоставляет определения с использованием следующих форм, таких как, но не ограничиваясь, буквальное описание, уравнения, таблицы, рисунки и легенды.

Стандарт также устанавливает словарь для разработки будущих стандартов для медицинских устройств с искусственным интеллектом.

IEEE P2807.2 Руководство по применению графов знаний для финансовых услуг

Этот стандарт определяет правила применения графов знаний для финансовых услуг. Стандарт определяет техническую структуру, рабочие процессы, рекомендации по внедрению и сценарии применения графов финансовых знаний.

IEEE P2817 Руководство по проверке автономных систем

Целью этого руководства является определение существующих передовых практик и предоставление наборов инструкций, которые определяют допустимые процессы проверки для ряда конфигураций автономных систем. Эти передовые методы применяются от компонентов и программного обеспечения самого низкого уровня до элементов обучения или принятия решений самого высокого уровня (в частности, включая проверку входных данных для любых алгоритмов обучения, таких как данные обучения). Рекомендации предназначены для включения как роботов, так и иммоботов, по отдельности и в группах, с упором в первую очередь на

системы, которые могут работать автономно, а не на автоматизированных или контролируемых роботов. Они также могут быть применимы к системам, которые напрямую не взаимодействуют с внешним миром (например, разведывательные сети).

IEEE P2830 Стандарт технической структуры и требований к общему машинному обучению

Этот стандарт определяет структуру и архитектуры для машинного обучения, в которых модель обучается с использованием зашифрованных данных, которые были агрегированы из нескольких источников и обрабатываются доверенным третьим лицом. Он определяет функциональные компоненты, рабочие процессы, требования безопасности, технические требования и протоколы.

IEEE P2841 Структура и процесс оценки глубокого обучения

В этом документе описываются передовые методы разработки и внедрения алгоритмов глубокого обучения, а также определяются структура и критерии для оценки надежности алгоритмов и качества получаемых программных систем.

IEEE P2842 Рекомендуемая практика для безопасных многосторонних вычислений

Этот стандарт обеспечивает техническую основу для безопасных многосторонних вычислений, включая определение: обзора безопасных многосторонних вычислений, технической основы безопасных многосторонних вычислений, уровней безопасности Безопасные многосторонние вычисления, варианты использования на основе безопасных многосторонних вычислений.

IEEE P2863 Рекомендуемая практика для организационного управления искусственным интеллектом

Эта рекомендуемая практика определяет критерии управления, такие как безопасность, прозрачность, подотчетность, ответственность и минимизация предвзятости, а также этапы процесса для эффективного внедрения, аудита эффективности, обучения и соответствия при разработке или использовании искусственный интеллект в организациях.

IEEE P2894 Руководство по архитектурной структуре для объяснимого искусственного интеллекта

Это руководство определяет архитектурную структуру, которая облегчает внедрение объяснимого искусственного интеллекта (ХАИ). Это руководство определяет структуру архитектуры и руководящие принципы применения для ХАИ, включая: 1) описание и определение объяснимого ИИ, 2) категоризацию объяснимых методов ИИ; 3) сценарии приложений, для которых необходимы объяснимые методы искусственного интеллекта, 4) оценки производительности ХАИ в реальных прикладных системах.

IEEE P3333.1.3 Стандарт для оценки визуального опыта на основе глубокого обучения с учетом человеческого фактора

Этот стандарт определяет основанные на глубоком обучении метрики анализа контента и оценки качества восприятия (QoE) для визуального контента, который является расширением стандарта для оценки качества восприятия (QoE) и визуального комфорта трехмерного (3D) содержимого на основе психофизических исследований (IEEE 3333.1.1 TM) и стандарта оценки качества восприятия трехмерного (3D) и сверхвысокого разрешения (UHD)) Содержание (IEEE 3333.1.2 TM). Объем охватывает следующее:

Модели глубокого обучения для оценки QoE (многослойные перцептроны, сверточные нейронные сети, глубокие генеративные модели)

Глубокие показатели визуального восприятия контента высокого разрешения (HD), UHD, 3D, расширенного динамического диапазона (HDR), виртуальной реальности (VR) и смешанной реальности (MR) * Глубокий клинический анализ (электроэнцефалограмма (ЭЭГ), электрокардиограмма (ЭКГ) , электроокулография (EOG) и т. д.) и психофизические (субъективные тесты и анкета симулятора болезни (SSQ)) данные для оценки QoE

Глубокая персонализированная оценка предпочтений визуального контента

Создание баз данных изображений и видео для целей тестирования производительности при необходимости

IEEE P3652.1 Руководство по архитектурной структуре и применению федеративного машинного обучения

Федеративное обучение определяет структуру машинного обучения, которая позволяет создавать коллективную модель из данных, распределенных между владельцами данных.

В этом руководстве представлен план использования данных и построения моделей в организациях с соблюдением применимых требований к конфиденциальности, безопасности и нормативным требованиям. Он определяет архитектурную структуру и руководящие принципы приложений для федеративного машинного обучения, включая: 1) описание и определение федеративного обучения, 2) типы федеративного обучения и сценарии приложений, к которым применяется каждый тип, 3) оценку производительности федеративного обучения и 4) соответствующие нормативные требования.

IEEE Std 7000 2021 Типовой процесс решения этических проблем при проектировании системы

В этом стандарте описывается подход к выявлению и анализу потенциальных этических проблем в системе или программном обеспечении с самого начала работы. Методы проектирования систем, основанные на ценностях, учитывают этические соображения на каждом этапе разработки, чтобы помочь избежать негативных непредвиденных последствий при одновременном увеличении инноваций.

IEEE P7001 Стандарты прозрачности автономных систем

Этот стандарт описывает измеримые, проверяемые уровни прозрачности, чтобы автономные системы можно было объективно оценить и определить уровни соответствия.

Основная проблема автономных систем (AS) заключается в том, что их работа должна быть прозрачной для широкого круга заинтересованных сторон по разным причинам.

Для разработчиков стандарт предоставит руководство по самооценке прозрачности во время разработки и предложит механизмы повышения прозрачности (например, необходимость безопасного хранения данных датчиков и внутреннего состояния, сравнимых с регистратором полетных данных или черным ящиком).

IEEE P7002 Стандарт процесса обеспечения конфиденциальности данных

Этот стандарт определяет, как управлять проблемами конфиденциальности для систем или программного обеспечения, которые собирают личные данные. Это будет сделано путем определения требований, которые охватывают корпоративные политики сбора данных и обеспечение качества. Он также включает вариант использования и модель данных для организаций, разрабатывающих приложения, содержащие личную информацию. Стандарт поможет разработчикам, предоставив способы выявления и измерения средств контроля конфиденциальности в их системах с использованием оценок воздействия на конфиденциальность.

IEEE P7003 Стандарт для рассмотрения алгоритмических предубеждений

этом стандарте описываются конкретные методологии, помогающие пользователям сертифицировать, как они работали над решением и устранением проблем, связанных с отрицательной предвзятостью при создании своих алгоритмов, где "отрицательная предвзятость" подразумевает использование чрезмерно субъективных или унифицированных данных. наборы или информация, заведомо несовместимые с законодательством в отношении определенных охраняемых характеристик; или со случаями предвзятого отношения к группам, не обязательно прямо защищенными законодательством, но иным образом снижающими благосостояние заинтересованных сторон или пользователей и для которых есть веские причины считать неуместными.

IEEE P7004 Стандарт для управления данными о детях и учениках

Стандарт определяет конкретные методологии, помогающие пользователям сертифицировать свой подход к доступу, сбору, хранению, использованию, совместному использованию и уничтожению данных о детях и учениках. Стандарт предоставляет конкретные показатели и критерии соответствия в отношении этих типов использования от надежных глобальных партнеров и того, как поставщики и образовательные учреждения могут их удовлетворить.

IEEE P7005 Стандарт прозрачного управления данными работодателя

Стандарт определяет конкретные методологии, помогающие работодателям сертифицировать свой подход к доступу, сбору, хранению, использованию, совместному использованию и уничтожению данных о сотрудниках. Стандарт предоставляет конкретные показатели и критерии соответствия в отношении этих

типов использования от надежных глобальных партнеров и того, как поставщики и работодатели могут их удовлетворить.

IEEE P7006 Стандарт для агента искусственного интеллекта (AI) персональных данных

В этом стандарте описываются технические элементы, необходимые для создания и предоставления доступа к персонализированному искусственному интеллекту (AI), который будет включать вводимые данные, обучение, этику, правила и ценности, контролируемые отдельными лицами.

IEEE P7007 Онтологический стандарт для робототехники и систем автоматизации, основанных на этических принципах

Стандарт устанавливает набор онтологий с различными уровнями абстракции, которые содержат концепции, определения и аксиомы, необходимые для разработки этически ориентированных методологий для проектирования роботов и систем автоматизации.

IEEE P7008 Стандарт этически обусловленного подталкивания для роботизированных, интеллектуальных и автономных систем

«Подталкивания», проявляемые роботизированными, интеллектуальными или автономными системами, определяются как явные или скрытые внушения или манипуляции, призванные повлиять на поведение или эмоции пользователя.

Этот стандарт устанавливает разграничение типичных подталкиваний (используемых в настоящее время или которые могут быть созданы). Он содержит концепции, функции и преимущества, необходимые для разработки и обеспечения этически ориентированных методологий проектирования роботизированных, интеллектуальных и автономных систем, которые их включают.

IEEE P7009 Стандарт отказоустойчивого проектирования автономных и полуавтономных систем.

Этот стандарт устанавливает практическую техническую базу конкретных методологий и инструментов для разработки, внедрения и использования

эффективных отказоустойчивых механизмов в автономных и полуавтономных системах. системы.

Стандарт включает (но не ограничивается ими): четкие процедуры измерения, тестирования и сертификации способности системы к безопасному отказу по шкале от слабого до сильного, а также инструкции по улучшению в случае неудовлетворительной работы.

IEEE 7010-2020 Рекомендуемая практика IEEE для оценки влияния автономных и интеллектуальных систем на благополучие человека

Рекомендуемая практика для оценки воздействия автономных и интеллектуальных систем (A / IS) на благополучие человека - это рекомендованная практика для измерения воздействия A / IS на людей. Общая цель IEEE P7010™ - поддержать результаты A / IS, оказывающие положительное влияние на благополучие людей.

Рекомендуемая практика основана на научно обоснованных индексах благополучия, используемых в настоящее время, и на процессе взаимодействия с заинтересованными сторонами. Целью рекомендуемой практики является руководство разработкой продукта, определение областей для улучшения, управление рисками, оценка производительности и выявление предполагаемых и непреднамеренных пользователей, использования и воздействия на человеческое благополучие A / IS.

IEEE P7011 Стандарт для процесса определения и оценки надежности источников новостей

Этот стандарт обеспечивает полуавтономные процессы, использующие стандарты для создания и поддержания рейтингов поставщиков новостей в целях повышения осведомленности общественности. Он стандартизирует процессы для выявления и оценки фактической точности новостных сообщений с целью составления рейтинга поставщиков новостей онлайн и части поставщиков мультимедийных новостей. Этот процесс будет использоваться для составления оценочных карточек правдивости с помощью многогранных подходов с использованием разных источников.

Стандарт определяет алгоритм, использующий программное обеспечение с открытым исходным кодом, и систему оценок в качестве методологии оценки надежности в качестве основного клиента в целях установления доверия и принятия.

IEEE P7012 Стандарт для машиночитаемых условий личной конфиденциальности

Стандарт определяет способ, которым предлагаются условия личной конфиденциальности, а также то, как они могут быть прочитаны и согласованы машинами.

IEEE P7014 Стандарт этических соображений при моделировании эмпатии в автономных и интеллектуальных системах

Этот стандарт определяет модель этических соображений и практик при разработке, создании и использовании эмпатической технологии, включая системы, способные определять, количественно оценивать и реагировать на них. или имитировать аффективные состояния, такие как эмоции и когнитивные состояния. Это включает в себя освещение «аффективных вычислений», «искусственного интеллекта эмоций» и связанных областей.

IEEE 1232.3-2014 Руководство IEEE по использованию обмена искусственным интеллектом и привязки сервисов ко всем тестовым средам (AI-ESTATE)

В этом руководстве представлены инструкции для разработчиков приложений, совместимых с IEEE Std 1232.

IEEE 1855-2016 Стандарт IEEE для языка нечеткой разметки

В этом стандарте представлен новый язык спецификаций, называемый языком нечеткой разметки (FML), использующий преимущества, предлагаемые спецификациями расширяемого языка разметки (XML) и соответствующими инструментами для моделирования система с нечеткой логикой, понятная человеку и не зависящая от оборудования.

IEEE 1873-2015 Стандарт IEEE для представления данных карты работа для навигации

В этом стандарте определено представление данных карты среды мобильного робота, выполняющего задачу навигации. Он предоставляет модели данных и форматы данных для двумерных (2D) метрических и топологических карт.

IEEE IC20-010 Маркировка данных кибербезопасности для автоматизации ИИ (одно- и мультимодальная)

Кибер-аналитики становятся узким местом при анализе постоянно растущих объемов данных. Автоматизация действий кибер-аналитиков с помощью ИИ может помочь сократить объем работы аналитиков и тем самым значительно сократить время до получения результата, записывать действия в базы знаний для обучения новых кибер-аналитиков и в целом открыть поле для новых возможностей. В результате состояние кибербезопасности улучшится. Предполагается, что эта группа соберет вместе заинтересованные стороны отрасли для достижения консенсуса по приоритетным вопросам для деятельности по стандартизации по этим темам и предоставит платформу для идейного лидерства IEEE в отрасли.

IEEE IC20-012 - Дорожная карта для разработки и внедрения стандартных ориентированных графов знаний

Это упражнение помогает организациям или пользователям, которые разрабатывают и применяют стандартизированные графы знаний, получить базовое представление о структуре и общем методе построения. Кроме того, это может помочь разработчикам графов знаний разработать общий интерфейс и следовать уточненным метрикам оценки. Кроме того, стандартные графы знаний можно интегрировать, внедрить и применять более просто и эффективно.

IEEE IC20-015 - Лаборатория Земли IEEE

Целью Лаборатории Земли IEEE является разработка Зеленого руководства по системам искусственного интеллекта (AIS), которое послужит прагматической дорожной картой для инженеров, корпоративных организаций и политиков по использованию инноваций AIS для эффективный переход к зеленой экономике. Мы достигнем этой цели, развивая и поддерживая глобальную сеть Living Labs, которая развертывает экологически согласованные ИСУ для эффективных, пригодных для жизни городов, низкоуглеродных, справедливых и устойчивых инфраструктур и

процветающих экосистем для сообществ, наиболее затронутых последствиями глобального потепления.

IEEE IC20-016 - Глобальная инициатива IEEE по этике расширенной реальности

Цель этой группы Industry Connections - продолжить и расширить существующие усилия Ассоциации стандартов IEEE, сосредоточенные на этических вопросах, связанных с расширенной реальностью, как указано в главе о расширенной реальности. этически согласованного дизайна, приглашая членов рабочих групп из нескольких рабочих групп по стандартам, ориентированных на дополненную и виртуальную реальность и пространственную сеть, а также дополнительных экспертов в предметной области из отрасли и политики для создания официальных документов, семинаров и PAR, связанных с этой работой, чтобы гарантировать их технологии переходят от «опасных» к «целенаправленным».

Целью данной программы отраслевых связей является усиление работы Ассоциации стандартов IEEE по биобезопасности и безопасности в соответствии с миссией IEEE «Развитие технологий для человечества».

IEEE IC20-027 - Ответственные инновации в области искусственного интеллекта и наук о жизни

Нигде потенциал искусственного интеллекта (AI) и автономных интеллектуальных систем (AIS) не является более очевидным, чем в здравоохранении и биологии человека, где все более сложные методы моделирования компьютерных данных привели к существенное улучшение нашей способности точно диагностировать и лечить заболевания, оценивать риски и оказывать помощь. Генетическая информация все чаще используется в алгоритмах искусственного интеллекта для выбора лечения и определения того, предоставляется ли лечение вообще. Преобразующее воздействие этих технологий и превращение наших биологических и геномных данных в товар окажет значительное влияние на будущий биологический континуум и геополитический порядок.

Литература

1. IEEE portfolio of AIS technology and impact standards and standards projects
<https://standards.ieee.org/initiatives/artificial-intelligence-systems/standards.html#p7000>
2. ПРАВОВОЕ РЕГУЛИРОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
Филипова И.А. http://www.lib.unn.ru/students/src/Prav_reg_I_I.pdf
3. Разработка рекомендации об этических аспектах искусственного интеллекта
ЮНЕСКО <https://ru.unesco.org/artificial-intelligence/ethics#recommendation>
4. Глобальная инициатива IEEE по этике автономных и интеллектуальных систем
<https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>
5. The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) Industry Connections Activity Initiation Document (ICAID) Version: 2.0, 23 September 2020
<https://standards.ieee.org/content/dam/ieee-standards/standards/web/governance/iccom/IC18-004-ECPAIS.pdf>
6. IEEE1872-owl - это спецификация OWL базовой онтологии для робототехники и автоматизации (CORA) и других онтологий в стандарте IEEE 1872-2015
<https://github.com/srfiorini/IEEE1872-owl>
7. Edson Prestes, Sandro Rama Fiorini, Joel Luis Carbonera, Core Ontology for Robotics and Automation // September 2014, Conference: Standardized Knowledge Representation and Ontologies for Robotics and Automation At: Chicago, Illinois, USA
https://www.researchgate.net/publication/273122687_Core_Ontology_for_Robotics_and_Automation