Научная статья
УДК 34:004.8:007.524:62-529.4
EDN: https://elibrary.ru/unsonv
DOI: https://doi.org/10.21202/jdtl.2023.2

Регулирование умных роботов и искусственного интеллекта в Европейском союзе

Кьяра Галлезе-Нобиле

Эйндховенский технологический университет
г. Эйндховен, Королевство Нидерландов;
Университет Триеста
г. Триест, Итальянская Республика

Ключевые слова

Европейский союз,
законодательство,
искусственный интеллект,
кибербезопасность,
машинное обучение,
право,
регулирование,
робот,
робототехника,
цифровые технологии

Аннотация

Цель: в последние годы в Европе стала очевидной необходимость создания нормативных актов в области робототехники и искусственного интеллекта. Европейский союз нуждается в стандартизации регулирования, которое обеспечило бы высокий уровень безопасности робототехнических систем и предотвратило потенциальные нарушения. Поэтому новое законодательство должно гарантировать, что именно производители несут ответственность за выявление «слепых зон» в этих системах, содержащих недоработки, а в случае обнаружения уязвимостей на поздних этапах – обновлять систему, даже если модель уже не находится на рынке. Цель данной статьи – предложить ряд возможных поправок к существующим положениям нормативных актов Евросоюза.

Методы: используя методологию Кестемонта для юридических исследований, автор анализирует и сравнивает тексты нормативных актов и увязывает их с техническими положениями об умных роботах, в результате выделяя критические элементы, требующие доработки.

Результаты: в статье предлагаются поправки к существующим проектам законов: по мнению автора, хотя закон об искусственном интеллекте и закон о киберустойчивости представляют собой первый шаг в этом направлении, их основные принципы недостаточно детализированы, а значит, не позволяют регулировать их практическое применение. Поэтому законодатели должны тщательно определить, в каких случаях модели, использующие обучение в течение всего жизненного цикла, могут быть допущены на рынок. Автор предлагает

© Галлезе-Нобиле К., 2023

Статья находится в открытом доступе и распространяется в соответствии с лицензией Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/deed.ru), позволяющей неограниченно использовать, распространять и воспроизводить материал при условии, что оригинальная работа упомянута с соблюдением правил цитирования.

увеличить сроки обязательного обновления систем, предусмотренные в существующих проектах законов, поскольку пятилетний срок не соответствует уровню риска, связанного с продуктами длительного пользования, такими как транспортные средства.

Научная новизна: автор рассматривает действующую нормативную базу в области систем искусственного интеллекта и устройств с цифровыми элементами, выделяет риски действующей правовой базы и вносит предложения по возможному усовершенствованию существующих проектов законов.

Практическая значимость: материалы статьи могут быть использованы для доработки существующих проектов закона об искусственном интеллекте и закона о киберустойчивости.

Для цитирования

Галлезе-Нобиле, К. (2023). Регулирование умных роботов и искусственного интеллекта в Европейском союзе. *Journal of Digital Technologies and Law*, *1*(1), 33–61. https://doi.org/10.21202/jdtl.2023.2

Содержание

Введение

Несмотря на заметные достижения Директивы о сетевой и информационной безопасности (Network and Information Security, NIS Directive) и предпринимаемые усилия по их углублению с помощью Директивы NIS II, проектов закона о киберустойчивости и закона об искусственном интеллекте, тема кибербезопасности в цифровую эпоху остается важнейшей для европейского рынка. Инструменты регулирования помогли нескольким государства – членам Евросоюза адаптировать свои институциональные, правовые и ценностные подходы к кибербезопасности, однако распространение

таких современных технологий, как системы искусственного интеллекта, а также такие беспрецедентные обстоятельства, как пандемия COVID-19, ускорили переход к цифровому сообществу через увеличение картины угроз и создание новых проблем, требующих решения на национальном и международном уровнях. Участились и стали более серьезными кибератаки, исходящие как изнутри, так и снаружи Евросоюза и часто создающие значительные риски для прав и свобод граждан[1]. Поскольку целями кибератак становятся все новые объекты оборудования и программного обеспечения, прогнозируется, что к 2021 г. ежегодный ущерб от киберпреступности в мире достигнет 5,5 трлн евро.

В Пояснительной записке к проекту закона о киберустойчивости показано, как недостаточный уровень понимания и низкая информированность пользователей мешают им выбирать продукты с адекватными возможностями киберзащиты и приводят к небезопасному потреблению[2]. Такие продукты подвержены двум серьезным рискам, которые могут привести к ущербу для пользователей и сообщества в целом: (1) недостаточной информационной безопасности, которая проявляется в постоянных нарушениях защиты, а также в неадекватном и нерегулярном появлении обновлений, необходимых для борьбы с ними, и (2) низкому уровню знаний и доступности информации для пользователей[3]. Кибератака на один продукт может затронуть целую организацию или цепь поставки, зачастую быстро распространяясь за пределы границ внутреннего рынка[4].

Реформирование правовых механизмов кибербезопасности находит поддержку среди предпринимателей и компетентных представителей власти. В ходе многочисленных консультаций они указывали, что Директиву NIS необходимо распространить на другие отрасли, интегрировать или оптимизировать дополнительные меры безопасности, а также упростить требования к отчетности[5]. Это говорит о том, что вопрос кибербезопасности воспринимается как важнейшая проблема, требующая решения на институциональном уровне.

В настоящей статье представлены некоторые результаты анализа проблем, связанных с сетевыми умными роботами в европейской системе права, и ряд предложений в этой сфере. В первой части представлен анализ определения искусственного интеллекта (далее – ИИ) и умных роботов, описаны инструменты регулирования в этой области. Затем приводятся некоторые соображения по поводу машинного обучения и принципа «черного ящика», выделены вопросы, требующие решения на уровне ЕС. Наконец, дается краткий анализ недостатков проекта закона о киберустойчивости.

---

1 Commission Staff Working document Delivering on the UN's Sustainable Development Goals – A comprehensive approach. https://sdgtoolkit.org/wp-content/uploads/2021/02/EU-staff_working_document-delivering_on_uns_sustainable_development_goals_en.pdf

2 European Commission. (2022). https://ireland.representation.ec.europa.eu/news-and-events/news/new-eu-cybersecurity-rules-ensure-more-secure-hardware-and-software-products-2022-09-15_en

3 Там же.

4 Там же.

5 Commission Staff Working document Delivering on the UN's Sustainable Development Goals – A comprehensive approach. https://sdgtoolkit.org/wp-content/uploads/2021/02/EU-staff_working_document-delivering_on_uns_sustainable_development_goals_en.pdf

1. Искусственный интеллект и роботы: определения

В последнее десятилетие использование систем на основе искусственного интеллекта, включая роботов, особенно использующих машинное обучение, достигло значительных масштабов и будет только расти в будущем (Wang & Siau, 2019). Среди примеров их современного применения – дроны (De Swarte et al., 2019), самоуправляемые автомобили (Rao & Frtunikj, 2018), медицинское оборудование (O'Sullivan et al., 2019) и средства диагностики (Nobile et al., 2019), бытовая техника (Yuniarthe, 2017), роботы для ухода за пожилыми (Simoens et al., 2016) и для совместной работы с людьми (Demir et al., 2019), переводчики (Bi, 2020), распознавание образов (Rundo et al., 2019), страхование (Lamberton et al., 2017), прогнозирование (например, юридических решений (Sourdin, 2018) или финансовых вопросов), аварийная связь (Raza et al., 2020), HR (Yano, 2017), управление дорожным движением (Aladin et al., 2019), управление электронной почтой (включая фильтры от спама (Das et al., 2015)), строительство (Chakkravarthy, 2019), логистика (Pandian, 2019), образование (Roll & Wylie, 2016) и множество других областей. Это быстрое развитие привело к тому, что уже несколько лет на повестке дня в Европейском союзе стоит вопрос правовой определенности в отношении новых технологий.

Уровень, достигнутый к настоящему времени в области ИИ, очень далек от того, что предсказывал Тьюринг («сильный ИИ»). Нынешний искусственный интеллект называют «узким ИИ» (Artificial Narrow Intelligence, ANI) (Kaplan & Haenlein, 2019), или «слабым ИИ» (Lu et al., 2017). Он состоит из систем, способных решать конкретные задачи на основе моделей и алгоритмов, созданных и отлаженных человеком, используя данные и параметры, произвольно выбранные создателем модели. Это означает, что он способен выполнять определенные задачи, иногда даже лучше, чем человек, но не может перенести эти навыки в другую область знаний. Следовательно, термин «интеллект» ошибочен.

Прежде всего, необходимо определить соответствующую юридическую терминологию, поскольку правовые определения часто не совпадают с терминологией ученых в области компьютерных наук (Smith, 2016), среди которых при этом также нет единогласия (Bekey, 2012).

Проект общеевропейского нормативного акта по ИИ (так называемый закон об ИИ, AI Act)[6] перечисляет технологии, которые на уровне ЕС считаются искусственным интеллектом; предлагаемое определение системы искусственного интеллекта является очень широким и включает технологии, которые известны уже не одно десятилетие: «<…> программное обеспечение, разработанное с использованием одной или более технологий и подходов, перечисленных в Приложении I, и способное для заданного набора определенных человеком целей порождать такие результаты, как содержание, предсказания, рекомендации или решения, влияющие на окружение, с которым оно взаимодействует». Фактически список в Приложении I содержит широкий спектр различных технологий: «a) подходы с точки зрения машинного обучения, включая контролируемое, неконтролируемое обучение и обучение с подкреплением, с использованием разнообразных методов, включая глубокое обучение; б) подходы на основе логики и знаний, включая представление знаний, индуктивное (логическое) программирование, базы

[6] С проектом можно ознакомиться по ссылке: https://ec.europa.eu/newsroom/dae/items/709090

знаний, механизмы логического вывода и дедукции, (символические) умозаключения и экспертные системы; в) статистические подходы, байесовское оценивание, методы поиска и оптимизации». Цель данного определения – включить множество различных моделей, способных оказывать влияние на человека, даже если они пока не относятся к ИИ. Это затруднит вывод приложения из-под регулирования путем простой замены определения модели так, чтобы она не относилась к ИИ. Однако данное определение все же подвергалось критике как слишком широкое[7].

Для термина «робот» сложно подобрать подходящее определение. По мнению экспертной группы High-Expert Group, «робототехнику можно определить как “ИИ, действующий в физическом мире”» (говорят также «воплощенный ИИ»). Робот – это физический механизм, которому приходится справляться с динамичностью, неопределенностью и сложностью физического мира. Такие функции, как восприятие, рассуждение, действие, обучение, а также способность взаимодействовать с другими системами, обычно интегрированы в архитектуру управления робототехнического комплекса. Кроме ИИ, строение и функционирование робота определяется также другими дисциплинами, такими как инженерная механика и теория управления. Примерами роботов являются роботизированные манипуляторы, беспилотные транспортные средства (например, машины, дроны, воздушные такси), гуманоидные роботы, роботы-пылесосы и др.

В исследовании «Европейские нормы гражданского права в сфере робототехники» (Nevejans, 2016), напротив, утверждается, что, поскольку научное сообщество еще не выработало единого определения роботов, нерешенными остаются и проблемы с терминами «умный робот» и «автономный робот» (Haselager, 2005). Некоторые авторы ссылаются на определение Richards и Smart (Richards & Smart, 2016), согласно которому робот – это «сконструированная система, выполняющая как физические, так и интеллектуальные действия, но не живая в биологическом смысле», однако при этом ведется настолько широкое обсуждение понятия «действия», что мы не можем рассматривать эту проблему здесь; ей будет посвящена отдельная статья.

Согласно решению Европейского парламента от 16 февраля 2017 г. с рекомендациями для Комиссии по нормам гражданского права в сфере робототехники (European Parliament, 2017), для признания в качестве умного робота с юридической точки зрения объект должен отвечать следующим условиям:

– получение автономии путем использования сенсоров и/или обмена данными с окружающей средой (совместимость), сравнения и анализа этих данных;

– самообучение на основе опыта и путем взаимодействия (необязательный критерий);

– по крайней мере, минимальное физическое воплощение;

– адаптация своего поведения и действий к окружающей среде;

– отсутствие жизни в биологическом смысле.

---

7 См., например, проект заявления Комитета по проблемам промышленности, науки и энергетики для Комитета по проблемам внутреннего рынка и защиты прав потребителей и Комитета гражданских свобод, правосудия и внутренних дел, доступный по ссылке: https://www.europarl.europa.eu/doceo/document/ITRE-PA-719801_EN.pdf , а также проект заявления Комитета по правовым проблемам для Комитета по проблемам внутреннего рынка и защиты прав потребителей и Комитета гражданских свобод, правосудия и внутренних дел.

Определения очень важны для совершенствования существующих и выработки новых нормативных актов, поскольку многие ситуации могут выпасть из сферы регулирования, если они не подпадают под юридическое определение. При этом отмечается, что любые попытки дать всеобъемлющее определение обречены на провал, так как роботизированные приложения чрезвычайно разнообразны (Bertolini, 2013). В нашей работе мы используем термин «устройство» в отношении инструментов и роботов на основе ИИ, независимо от их типа, и термины «программное обеспечение» или «модель» в отношении систем ИИ, включая установленные на указанных устройствах.

Мы считаем, что с юридической точки зрения нет существенной разницы между механизмом, который можно назвать роботом, и управляемым искусственным интеллектом – устройством, которое нельзя в строгом смысле назвать роботом, но которое тем не менее является умным и может оказывать влияние на окружающую среду. Иногда граница между ними довольно подвижна. Например, являются ли роботами умная стиральная машина или робот-пылесос? Беспилотный автомобиль? Домашний помощник или мобильный телефон, управляющий дверями, светом, вентиляцией в доме, воротами гаража и т. д.? А что насчет 4D-принтера? Если это не роботы, то где должна проходить юридическая граница?

Если следовать определению, данному в вышеуказанной резолюции, то даже игрушечный утенок (отсутствие жизни и физическое воплощение), двигающийся по комнате и меняющий направление при блокировке колес о препятствие (адаптация к окружающей среде), имеющий таймер запуска в действие и исполняющий различные мелодии при перемене освещенности (получение автономии путем использования сенсоров), будет считаться умным роботом, даже если он состоит лишь из электрических цепей и механических деталей без какого-либо программного обеспечения или модели искусственного интеллекта.

Таким образом, данное определение довольно расплывчато и недостаточно для полного описания умного робота; оно включает механизмы, не обладающие никаким ИИ.

Предлагалось также считать ключевым аспектом робота его способность использовать программное обеспечение при выполнении конкретных задач (Leenes et al., 2017), однако, по нашему мнению, это неверно (так можно определить и компьютер!); фактически ни автономное выполнение задач (Santosuosso et al., 2012), ни наличие программного обеспечения не позволяют отграничить роботов от других механизмов. Например, достаточно вспомнить, что роботы могут создаваться с использованием таких физических компонентов, как мемристоры, которые выполняют ту же функцию, что и биологические синапсы в нейроморфных устройствах, не требуя при этом никакого программного обеспечения (Ames et al., 2012).

Программное обеспечение может быть загружено в робота, но может и существовать вне его, как в случае виртуальных роботов, которые подлежат регулированию как программное обеспечение. Более того, физические компоненты робота могут быть не связаны с конкретным программным обеспечением, которое при этом может продаваться отдельно. Код, управляющий роботом, может существовать как единое целое, но зачастую он составляется из множества взаимодействующих между собой модулей, где каждый компонент выполняет отдельную задачу.

В сфере ИИ важно различать концепции программного обеспечения, нейроморфных устройств и схем, а также программируемую пользователем вентильную

матрицу (ППВМ, FPGA). Различные технологии могут комбинироваться между собой, как в чипе Spikey[8]. При создании умного робота могут быть использованы любые решения, но правовые последствия для производителей и программистов будут различными.

Напротив, нейроморфные устройства (Ielmini & Ambrogio, 2019) и схемы (Pan et al., 2020) – это физические объекты, функционирующие только в конкретном роботе, в который они встроены; они разрабатываются для выполнения нейровычислений, но также используются для симуляции динамики биологических нейросетей (Papetti et al., 2020). Технология ППВМ (Botros & Abdul-Aziz, 1994) является промежуточной по отношению к двум вышеописанным решениям: это интегрированная физическая схема, которая может быть конфигурирована конечным пользователем или программистом.

Все эти решения могут использоваться при создании роботов, а в будущем, возможно, и функционирующего физического искусственного мозга.

Программируемость также не является существенным признаком робота: например, в принципе, возможно создать робота на магнитных модулях, который будет двигаться, самостоятельно выполнять действия и общаться с другими роботами световыми сигналами, при этом его создатель не будет знать, как он себя поведет, как в концепции М-куба (Romanishin et al., 2013).

Мы считаем, что программное обеспечение с использованием ИИ, которое может устанавливаться на любые устройства, должно регулироваться единым законодательством, независимо от того, установлено оно на устройство, имеющее в настоящий момент статус робота или нет; и наоборот, роботы, не относящиеся к категории умных роботов, особенно если они не связаны с Интернетом (как, например, промышленные роботы на заводском конвейере), должны относиться к сфере обычного законодательства о дефектной продукции и других норм, применимых к машинам. Различие следует проводить, руководствуясь влиянием на человека и общество, а не наличием физического объекта.

Основная причина, по которой необходимо объединить роботов и программное обеспечение на основе искусственного интеллекта в рамках одной дисциплины, состоит в том, что при их разделении будет возможно продавать программное обеспечение (далее – ПО) и устройство отдельно, не подпадая при этом под регулирование об умных роботах. Тогда любой пользователь сможет установить другое ПО и создать умного робота, запрограммировав его на выполнение любой деятельности. Следует отметить, что, по доминирующему мнению, не может считаться роботом устройство, дистанционно управляемое человеком, если человек должен присутствовать для выполнения определенных действий или контроля над устройством (Ebers & Navas, 2020; Funkhouser, 2013). Следовательно, физическое устройство без установленного на нем программного обеспечения (или доступного ему, как в случае выполнения ПО на облаке) также не может считаться роботом.

Следует также упомянуть, что робот, используемый в промышленном производстве, считается устройством, а значит, подпадает под Директиву 2006/42/ЕС Европейского парламента и Совета от 17 мая 2006 г. и под Директиву 95/16/ЕС32.

---

8 Дополнительную информацию см.: https://www.kip.uni-heidelberg.de/vision/previous-projects/facets/neuromorphic-hardware/single-chip-system/spikey

Европейский проект закона об ИИ фактически ссылается на Директиву об устройствах в аспекте включения умных роботов в общее регулирование ИИ, тем самым объединяя физических и виртуальных роботов.

2. Регулирование искусственного интеллекта и умных роботов в ЕС: технические регламенты

Когда тема искусственного интеллекта начала все более активно обсуждаться среди ученых в области компьютерных наук всего мира, европейские правоведы и законотворцы задумались о правовых аспектах ИИ и умных роботов, сосредоточившись в основном на проблемах ответственности производителя, защиты прав потребителей и технического регулирования. В юридической доктрине была признана насущная необходимость принять сбалансированные отраслевые технические регламенты, которые отвечали бы потребностям самых разнообразных технологий роботостроения (Amidei, 2017).

При этом во многих случаях проблемы, вызванные роботами или устройствами с ИИ, рассматривались постфактум. Однако важно также понимать, как предотвращать аварии, особенно в области кибербезопасности. Даже если невозможно избежать проблемы в целом, важно анализировать ее с точки зрения предотвращения риска (такой подход используется в Директиве об устройствах), и новый проект закона об ИИ движется в этом направлении.

Самая важная инновация проекта закона об искусственном интеллекте – это положение о четырех категориях риска, связанного с системами ИИ, направленное на защиту основных прав граждан[9].

[9] Фактически в пояснительной записке к проекту отмечается: «Использование ИИ с его отличительными признаками (такими как непрозрачность, сложность, зависимость от данных, автономность действий) может отрицательно повлиять на количество фундаментальных прав, закрепленных в Европейской хартии по правам человека (“Хартия”). Данный проект нацелен на обеспечение высокого уровня защиты данных фундаментальных прав и работу с различными источниками рисков путем четко определенного подхода. Устанавливая ряд требований к надежному ИИ и пропорциональные обязательствам всех участников цепочки создания ценности, данный проект повысит и усилит защиту прав, закрепленных в Хартии: право на человеческое достоинство (ст. 1), уважение к частной жизни и защиту персональных данных (ст. 7 и 8), отсутствие дискриминации (ст. 21) и равенство между мужчиной и женщиной (ст. 23). Он нацелен на предотвращение нарушения прав в области свободы слова (ст. 11) и свободы собраний (ст. 12), на обеспечение защиты права на эффективное восстановление в правах и справедливое правосудие, прав на юридическую защиту и презумпцию невиновности (ст. 47 и 48), а также общего принципа компетентного руководства. Кроме того, при применении в определенных областях данный проект положительно повлияет на права ряда особых групп, например, права рабочих и справедливые условия работы (ст. 31), высокий уровень защиты прав потребителей (ст. 28), права ребенка (ст. 24) и интеграцию лиц с ограниченными возможностями (ст. 26). Сюда относятся также права на высокий уровень защиты и улучшение качества окружающей среды (ст. 37), в том числе в отношении здоровья и безопасности людей. Обязанности по априорному тестированию, управлению рисками и пользовательскому надзору будут также способствовать уважению других фундаментальных прав путем минимизации риска принятия ошибочных или предвзятых решений с помощью ИИ в таких критических областях, как образование и воспитание, занятость, важные услуги, правоохранительная и судебная деятельность. Если ущемление фундаментальных прав все же произойдет, пострадавшие лица смогут получить эффективное восстановление путем обеспечения прозрачности и отслеживаемости систем ИИ в сочетании со строгим последующим контролем».

Категории риска относятся к степени (интенсивности и масштабу) риска для безопасности или основных прав граждан и классифицируются по трем группам (четвертая группа – отсутствие риска):

– (i) неприемлемый риск,
– (ii) высокий риск,
– (iii) низкий или минимальный риск.

Данная классификация берет начало в законодательстве о безопасности товара и основана на предусмотренном назначении и условиях использования системы ИИ, а не только ее конкретной функции[10]. В проекте также приводится список запрещенных систем искусственного интеллекта, попадающих в первую категорию риска.

В рамках такого подхода умные роботы относятся к категории высокого риска (фактически, как уже говорилось, это ссылка на Директиву об устройствах). Такая классификация приводит к выработке новых требований, и программисты/производители должны применять следующее:

– систему управления риском;
– систему распоряжения данными;
– меры по обеспечению прозрачности;
– меры пользовательского надзора;
– меры по обеспечению точности, последовательности кибербезопасности;
– систему управления качеством.

После появления проекта закона об ИИ Еврокомиссия опубликовала новый проект относительно кибербезопасности: так называемый закон о киберустойчивости[11], нацеленный на обеспечение: 1) повышенной безопасности изделий с цифровыми элементами в течение всего их жизненного цикла; 2) гармонизации структуры кибербезопасности в Европе; 3) повышенной прозрачности свойств безопасности таких изделий; 4) повышенной безопасности таких изделий для организаций и потребителей.

В дополнение к новому подходу были опубликованы еще два проекта: Директива об ответственности в сфере ИИ и новая редакция Директивы об ответственности производителя.

В следующих разделах мы покажем, как эти требования могут быть реализованы и расширены за счет новых положений на этапе рассмотрения нового законодательства.

3. Законодательные решения для предотвращения кибератак

Следует отметить, что с развитием технологий обнаруживается все больше ошибок и дефектов в старом программном обеспечении (Buchanan, 2016), которые приводят к уязвимости перед киберугрозами (Ozkan & Bulkan, 2019) и требуют усовершенствования средств защиты. Среди примеров недавних атак можно упомянуть первый зарегистрированный случай смерти от кибератаки в сентябре 2020 г., когда из-за

[10] В зависимости от правовой системы конкретного государства, квалификация высокого риска может повлечь иные последствия, помимо ответственности, см., например, ст. 2050 Гражданского кодекса Италии.

[11] Cyber-resilience Act. https://ireland.representation.ec.europa.eu/news-and-events/news/new-eu-cybersecurity-rules-ensure-more-secure-hardware-and-software-products-2022-09-15_en

взлома 30 компьютеров через программу-вымогатель в больнице Дюссельдорфа (Германия) медицинская помощь не была оказана вовремя[12]; взлом умных камер наблюдения в 2019 г., результатом которого стал коллективный судебный иск в США[13]; обнаружение уязвимостей кардиоимплантов в больнице св. Иуды, с помощью которых хакеры могли разрядить батарею, вызвать неправильный ритм или сердечный приступ[14]; бот-сеть Мираи (Zhang et al., 2020), с 2016 г. атаковавшая ряд устройств интернета вещей. Угрозы системам ИИ могут быть физическими: в литературе описаны примеры, показывающие, как цветовые поля[15] могут мешать работе нейронной сети, используемой беспилотными машинами (Ranjan et al., 2019), что приводит к неверному распознаванию или классификации сигналов и может вызвать аварии.

Очевидно, что борьба с кибератаками приобретает чрезвычайную важность, когда речь идет о фундаментальных правах, например, в сфере здравоохранения, которая особенно часто становится целью (Luna et al., 2016), или беспилотных машин, военной техники, спутников и других объектов, используемых для государственной обороны.

Производители и программисты постоянно обновляют программное обеспечение и устройства, пока не перейдут на новый продукт или новую версию существующего продукта. Через некоторое время они перестают разрабатывать защитные патчи и обновления для устаревших устройств, поскольку становится слишком дорого поддерживать тысячи продуктов на рынке, которые будут куплены лишь очень небольшим числом покупателей, или разрабатывать обновления для тех продуктов, которые уже ушли с рынка.

Чтобы программное обеспечение или операционные системы работали на определенных устройствах, производители последних вынуждены писать драйверы специально для этих устройств; поскольку такое ПО зачастую не является открытым, обновить эти устройства может только их производитель. Это означает, что устаревшие устройства, операционные системы и программное обеспечение в определенный момент остаются защищенными (например, Windows XP используется почти на 4 % компьютеров в мире, хотя уже не поддерживается компанией Microsoft), тем самым становясь легкой мишенью для атак. В некоторых случаях производители хотели бы обновить свои устройства, но не могут сделать этого, потому что осталось уже очень мало людей, которые понимают и могут правильно использовать древний или экзотический язык программирования (например, COBOL, который используется в банках; ADA – в военных приложениях, системах контроля авиатранспорта, коммерческих ракетах, спутниках, железнодорожном и скоростном транспорте; FORTRAN – в научных вычислениях), либо потому, что исходный код слишком длинный и сложный, чтобы его можно было изменить.

---

12 Подробнее см. www.bbc.com/news/technology-54204356. Программа-вымогатель – это вредоносная программа, блокирующая доступ к данным пользователя и иногда угрожающая раскрыть их третьим лицам, если жертва не заплатит определенную сумму.

13 С полным текстом можно ознакомиться на: https://www.courtlistener.com/docket/16630199/1/orange-v-ring-llc

14 Об этом сообщало агентство CNN: https://money.cnn.com/2017/01/09/technology/fda-st-jude-cardiac-hack

15 Цветовое поле – это особый вид вредоносных патчей, технология машинного обучения, которая с помощью ложной информации пытается обмануть нейросети и вызвать их неправильное функционирование.

Если существующая система останется такой же, вполне вероятно, что производители не будут обновлять умных роботов в течение всего их жизненного цикла, поскольку по действующему законодательству производители не несут ответственности по истечении трех лет после обнаружения дефекта потребителем в рамках 10-летнего периода, тогда как проект закона о киберустойчивости предлагает защиту лишь в течение пяти лет. Однако некоторые роботы могут прослужить более 10 лет (например, автомобили только приобретают ценность через 30 лет; см., например, итальянское понятие «auto d'epoca» – «автомобиль эпохи»), а значит, останутся без защиты через несколько лет эксплуатации.

Техническое регулирование в этих областях может быть достигнуто тремя путями: запретом устаревших технологий, которые представляют риск для фундаментальных прав человека; принуждением производителей к обновлению устаревших устройств и программного обеспечения; комбинацией двух этих решений. Каждый вариант имеет свои плюсы и минусы.

Наконец, при создании нового законодательства по умным роботам необходимо учитывать свойства различных типов роботов, чтобы сформулировать конкретные положения для каждого из них. Некоторые роботы создаются самостоятельно путем взаимодействия между объектами (например, модульные роботы создают новый объект или в будущем роботы будут создавать роботов), поэтому человеку сложно предвидеть, как они будут собраны. Роботов такого типа следует отнести к категории высокого риска из-за непредсказуемости их поведения. Некоторые роботы могут передвигаться или влиять на окружающую среду, потенциально представляя риск нанесения вреда людям, тогда как другие не могут этого делать (например, светофоры). Некоторые действуют без надзора человека (например, автоматическая торговля на бирже), представляя значительный риск, а другие напрямую управляются человеком или иными внешними силами (такими как свет, звуки, химикаты, магниты, электричество и даже животные); некоторые выполняют действия и принимают «решения», оказывающие значительное влияние на определенные права, а другие имеют лишь ограниченное влияние. Наконец, некоторые связаны с Интернетом или другой сетью (интернет вещей) и потенциально подвержены кибератакам, а другие нет. Все эти аспекты необходимо тщательно рассматривать при применении технического регулирования.

Новый проект закона об искусственном интеллекте не касается вопроса регулирования умных роботов. Фактически положение о кибербезопасности звучит расплывчато: «Статья 15. Достоверность, устойчивость и кибербезопасность. 1. Системы ИИ, относящиеся к категории высокого риска, должны разрабатываться и создаваться таким образом, чтобы достичь, с учетом их предполагаемого назначения, надлежащего уровня достоверности, устойчивости и кибербезопасности, и последовательно действовать соответствующим образом в течение всего их жизненного цикла. 2. Уровни достоверности и соответствующей достоверности измерений в системах ИИ категории высокого риска должны быть указаны в прилагаемых пользовательских инструкциях. 3. Системы ИИ категории высокого риска должны обладать устойчивостью к отказам в отношении ошибок, неточностей и непоследовательностей, которые могут возникнуть внутри системы или в среде, где работает система, в особенности при ее взаимодействии с физическими лицами или другими системами. Устойчивость систем ИИ категории высокого риска может быть достигнута через технически избыточные решения, в том числе системы дублирования или

защиты от отказа. Системы ИИ категории высокого риска, которые продолжают обучаться после выхода на рынок или ввода в эксплуатацию, должны разрабатываться так, чтобы были приняты соответствующие меры защиты против возможных искажений в результате использования искаженных выходных данных на входе будущих операций («петля обратной связи»). 4. Системы ИИ категории высокого риска должны обладать устойчивостью к отказам в отношении попыток неуполномоченных третьих лиц изменить их использование или функционирование через уязвимости системы. Технические решения, направленные на обеспечение кибербезопасности систем ИИ категории высокого риска, должны быть адекватны соответствующим обстоятельствам и рискам. Технические решения для борьбы с конкретными уязвимостями ИИ должны включать, где это необходимо, меры для предотвращения и контролирования атак, направленных на манипулирование обучающими базами данных (“заражение данных”), на ввод данных с целью заставить модель ошибаться (“вредные примеры”) и на недостатки модели».

В следующем разделе мы сфокусируем внимание на машинном обучении (контролируемом и неконтролируемом) и черных ящиках, так как они имеют особые свойства.

3.1. Машинное обучение и черные ящики

Благодаря научно-фантастической литературе мы привыкли к образу умного самообучающегося робота, однако в действительности лишь небольшая часть роботов запрограммирована для обучения на своем опыте. Некоторые роботы выглядят как умные роботы, так как выполняют сложные задачи и взаимодействуют с окружающей средой, но на самом деле они лишь запрограммированы с помощью простых алгоритмов.

С юридической точки зрения модели на основе машинного обучения (особенно глубокого обучения и непрерывного обучения) заслуживают пристального внимания благодаря своим характеристикам, поэтому мы обсудим их в отдельном подразделе.

3.1.1. Нормативные акты и стандарты безопасности для моделей на основе машинного обучения

Как известно, невозможно априори исключить наличие отклонений или дефектов (Naur & Randell, 1968), которые могут повлиять на поведение или решения, принимаемые устройством на основе искусственного интеллекта. Фактически модель создается человеком, который, разумеется, может ошибаться. Это означает, что программист никогда не сможет гарантировать, что модель поведет себя в точности так, как было запрограммировано[16], но также это означает, что, меняя гиперпараметры и выбирая данные так, как удобно (или политически выгодно), можно при помощи того же инструмента получить иные результаты.

Используемые данные могут сами по себе представлять ту же проблему: они могут быть неверными, ошибочными, неполными или просто слишком скудными, чтобы иметь какой-то смысл[17], поэтому результат будет ненадежным; более того, даже

[16] Даже проведя множество тестов, нельзя исключить наличие багов. По словам Е. В. Дейкстра, известного специалиста в области компьютерных наук, обладателя премии Тьюринга, «тестирование показывает присутствие, а не отсутствие багов» (Dijkstra,1982).

[17] Фактически машинное обучение модели часто требует использования тысяч примеров.

если данные корректны, искусственный интеллект может сделать выводы, предвзятые по отношению к определенным группам[18], на основе процессов, происходящих из-за неравенства, существующего в обществе, что в конечном итоге может нарушить соотношение между причиной и следствием, как показали некоторые широко известные эпизоды (Falletti, 2020).

Учитывая все вышесказанное, необходимо определить некоторые существенные риски и стандарты, которым должна отвечать модель искусственного интеллекта перед выпуском ее на рынок. По нашему мнению, ИИ следует моделировать по следующим общим критериям, часть из которых уже входит в Перечень для оценки надежности искусственного интеллекта (Assessment List for Trustworthy Artificial Intelligence, ALTAI):

– он должен быть понятным (в литературе чаще всего используются термины «прозрачность», «интерпретируемость», «объяснимость», «определимость» (Chakraborti et al., 2021; Gilpin et al., 2018; Holzinger et al., 2019), но мы считаем, что корректным термином является «интерпретируемость»), т. е. должны представляться прозрачными способы, которыми модель приходит к тем или иным решениям, так что всегда можно понять причины, по которым она делает выбор, а результат и обоснование решения модели всегда должны быть понятны людям;

– он должен соблюдать тайну частной жизни по умолчанию (Hildebrandt, 2019), т. е. модель должна отвечать требованиям европейского Общего регламента по защите данных в отношении данных пользователя и третьих лиц, а согласие на их использование должно быть информированным[19];

– он должен быть этичным (Greene et al., 2019), т. е. должно быть установлено правило избегания любого дискриминирующего поведения, которое может возникнуть, когда робот будет выпущен в общество, даже если такое поведение еще не предусмотрено действующим законом[20];

– он должен быть непредвзятым (Dietterich & Kong, 1995), т. е. обучение ИИ должно проходить под контролем экспертов, чтобы устранить любую возможную предвзятость, одновременно добиваясь надежности и точности и соблюдения применимых законов и норм[21];

– он должен быть современным и безопасным (Barreno et al., 2010), т. е. на регулярной основе должны применяться обновления и патчи, проводиться проверки,

---

18 Правовые аспекты машинного обучения в отношении возможной предвзятости и ее влияния на уязвимые группы будут рассмотрены в отдельной статье, так как эта тема слишком обширна для разбора здесь.

19 Хотя этот элемент уже присутствует в европейском законодательстве, граждане ЕС могут пострадать, если их данные будут собраны системами искусственного интеллекта за пределами Евросоюза.

20 Следует отметить, что в настоящей работе мы не исследуем исключительно права, защищаемые законом в каждой правовой системе, так как тот факт, что какое-либо право еще не признано государством, не означает, что оно не существует или что оно не будет в дальнейшем введено в эту систему. Некоторые проблемы, возникающие при использовании ИИ, проистекают не от багов и других видов ошибок программирования, а от политических решений или ненамеренной предвзятости человека.

21 Важно заметить, что мы используем слово «предвзятость» в понимании юридической науки, которое отличается от понимания компьютерных наук; фактически с точки зрения программиста можно показать, что действительно непредвзятой модели не может существовать. В настоящей работе мы рассматриваем только такую предвзятость, которая с правовой и этической точек зрения может повлечь дискриминацию или нарушение прав личности.

а механическая часть должна также проходить регулярный осмотр и ремонт в течение всего жизненного цикла робота;

– он должен соответствовать четким стандартам (O'Sullivan et al., 2019), т. е. в ЕС должен применяться ряд норм в отношении минимальных стандартов, которым должен соответствовать каждый робот в своей области функционирования, не только с точки зрения качества, но и с позиций безопасности. Это требование также содержится в проекте закона о киберустойчивости.

Проект закона об ИИ содержит некоторые из этих пунктов, но не рассматривает вопросы объяснимости и интерпретируемости в важнейших областях. Фактически требования в тексте проекта, в частности в Статье 10, относятся не к объяснимости и интерпретируемости, а скорее к информации, которая должна быть представлена в технической документации о функционировании системы ИИ.

Мы считаем, что недальновидно использовать модель черного ящика, по крайней мере, в случаях, когда затрагиваются здоровье и другие фундаментальные права, но также и во всех случаях, когда экономический ущерб может быть разрушительным для благосостояния пользователей (например, настолько крупный ущерб, что пользователи лишаются дома из-за рискованных инвестиций, сделанных роботом). Более того, мы считаем, что, когда речь идет о персональных данных, модели черного ящика не отвечают принципам Общего регламента по защите данных[22].

Хотя Европейский союз может ввести строгие регламенты и стандарты безопасности в дополнение к существующим регламентам (например, в медицинской области – нормы для медицинского оборудования), согласно их внутренней природе, однако модели с непрерывным обучением в течение всего жизненного цикла должны строго контролироваться во всех областях, где может возникнуть потенциальный вред для человека, даже если это создает негативные последствия для рынка. Защита пользователей (особенно пациентов, несовершеннолетних и представителей уязвимых групп населения) должна быть важнее, чем обеспечение разработки новых технологий в промышленном секторе, поскольку последнее все равно достигается научными исследованиями и государственными инвестициями. Таким образом, пресловутый сдерживающий эффект регламентов безопасности (как и правил ответственности производителя) на технологический прогресс не может считаться достаточным аргументом влияния на законодателей.

Что касается военных и других потенциально опасных устройств с ИИ, государство должно обеспечить специфичный и строгий контроль над ними, как это делается сейчас по отношению к опасным продуктам и некоторым суперкомпьютерам, для использования которых необходима лицензия, а перед утилизацией проводятся проверки данных. Такие ИИ-устройства никогда не должны выпускаться для широкой публики, не говоря уже об открытом исходном коде. Например, если бы умные роботы могли пользоваться оружием или иным образом убивать людей, а кто-то построил бы такого робота благодаря открытому исходному коду и доступу к 3D- и 4D-принтерам, то его можно было бы легко использовать для гражданской

[22] Проблема соответствия моделей машинного обучения принципам Общего регламента по защите данных выходит за рамки данной статьи и будет исследована в отдельной работе в ближайшие месяцы. Вопрос интерпретируемости и объяснимости более подробно разбирается в главе нашей выходящей книги «Правовые аспекты ИИ в области биомедицины. Роль интерпретируемых моделей».

войны или совершения террористических атак, особенно учитывая, что напечатанное пластиковое оружие не опознается металлодетекторами[23] (Falletti, 2022). Такой риск должен всерьез рассматриваться государственными органами тех стран, где законодательство по контролю над оружием менее строгое.

Что касается обязательных проверок, обновления и применения патчей безопасности для минимизации риска кибератак, именно государство должно вводить строгие проверки и аудит. Для потенциально опасных устройств, таких как хирургические роботы, должна вводиться обязательная регистрация на уровне ЕС.

Учитывая все вышесказанное, в том, что касается машинного обучения, мы считаем необходимым введение специфического регулирования на общеевропейском уровне. Существующие нормы не являются адекватными для регулирования всего комплекса вопросов, которые могут возникнуть после выхода таких технологий на рынок. В частности, учитывая неоднородный характер технологий на основе машинного обучения, а также отраслевой характер их конкретного применения, представляется невозможным создание всеобъемлющего законодательства, однако целесообразно регулировать каждую отрасль по отдельности. При этом можно подчеркнуть специфику каждой отрасли (например, диагностические, хирургические, автомобильные, образовательные устройства и т. д.) и соответственно адаптировать под нее регламенты.

Как было отмечено еще в работе Amidei, затягивание принятия новых регламентов при наличии запроса рынка может привести к огромным убыткам и рискам, когда под влиянием различных факторов придется принимать несправедливые нормы и следовать им, а также к риску неравноценных норм в государствах – членах ЕС, если они будут приняты в разное время (Amidei, 2017).

К сожалению, проекты европейских законов об искусственном интеллекте и о киберустойчивости не дают удовлетворительного решения по регулированию умных роботов, поскольку не учитывают множество разнообразных аспектов в этой области.

3.2. Правила и стандарты защиты для производителей и программистов

Если производители будут вынуждены обновлять свое программное обеспечение при выявлении новых угроз, то нагрузка ляжет на компанию, а не на миллионы граждан; однако это не всегда возможно. Например, такой подход заставит компанию держать сотрудника, потому что он один знает устаревший язык программирования, или инвестировать в создание патча безопасности, используемого только в одном типе устройств. Одна из возможных исправительных мер – обязательное профессиональное обучение сотрудников (подобное тому, что требуется в Италии для регулируемых профессий согласно Декрету Президента республики № 137/2012) в области устаревших или экзотических языков программирования и обновления устаревших технологий. Однако в некоторых случаях такое решение невозможно, так как код слишком сложный (например, в области авиационной бортовой радиоэлектронной

[23] Фактически даже в тех странах, где контроль над оружием менее строг, законодатели пытаются установить запрет на самостоятельную печать оружия; этот вопрос поясняется в прессе и в нашей выходящей статье «Этические и юридические ограничения на распространение самостоятельно произведенного оружия»: https://www.markey.senate.gov/news/press-releases/senator-markey-rep-meng-lead-colleagues-in-urging-biden-to-roll-back-trumps-deregulation-of-3d-printed-ghost-guns

аппаратуры: полный спектр функционирования самолета-истребителя F-35 требует исходного кода длиной 24 млн строк[24]). В других случаях закон мог бы заставить производителя заменить там, где это возможно, устаревший код на новый, тем самым устранив часть проблемы устаревания; фактически решения отказаться от конкретной технологии и прекратить выпуск обновлений часто оставляются на усмотрение производителя, даже если негативный эффект от нарушения безопасности ляжет на потребителей. То же верно и для тех решений по безопасности, которые могли бы предотвратить большое количество кибератак, но не являются обязательными для устройств из интернета вещей; например, это обязанность обновлять программное обеспечение устройства, пароли, физические части устройства, менять заложенное по умолчанию имя пользователя и пароль, требование использовать уникальный пароль для каждого устройства из интернета вещей.

Реализация подобного подхода могла бы изменить нашу экономику, возможно, замедлив технологический прогресс, но оказав положительный эффект на устойчивость развития за счет замедления устаревания умных продуктов.

Недавний пример, ярко показывающий необходимость нормативно-правового решения, которое заставило бы программистов обновлять программное обеспечение, – это уязвимость CVE-2018-13379 программного обеспечения компании Fortinet, о которой компании стало известно в 2018 г.[25], но обновление появилось только в мае прошлого года. Было обнаружено, что если задействован протокол безопасных соединений (SSL) сервиса виртуальной частной сети (VPN), то хакеры могут получить полномочия пользователей через обход каталога и удаленно скачать файлы системы FortiOS без авторизации. В Сеть был немедленно выложен список из более 49 тысяч VPN-адресов Fortinet FortiGate, которые могли быть атакованы хакерами.

Что касается кибербезопасности устройств, первый шаг в верном направлении был сделан установлением общей системы сертификации в Регламенте (Евросоюза) 2019/881 Европейского парламента и Совета от 17 апреля 2019 г. относительно Европейского агентства по сетевой и информационной безопасности (European Network and Information Security Agency, ENISA), хотя в ст. 56 этого Регламента и говорится, что «сертификация кибербезопасности является добровольной». Установление обязательной сертификации на общеевропейском уровне и системы наказаний за ее неисполнение явилось бы мощным инструментом в борьбе с киберугрозами.

Другим возможным решением стало бы принуждение производителей к использованию открытого программного обеспечения, созданного и применяемого в ЕС, с перспективой правдивого предоставления информации. Этот подход благоприятен по многим причинам. Во-первых, он позволяет легко и быстро выявить дефекты и создать патчи безопасности, так как множество программистов в разных странах Евросоюза (обладающие различными знаниями) будут работать одновременно. Кроме того, если программное обеспечение создается не производителями оборудования, то они не несут ответственности за дефекты в нем, даже в рамках законодательства об ответственности производителя, однако такую ответственность могут нести страны – члены Евросоюза.

---

24 Объяснение этой проблемы см.: https://spectrum.ieee.org/f35-program-continues-to-struggle-with-software

25 Эта уязвимость более подробно описывается в блоге компании, которая ее обнаружила: https://blog.orange.tw/2019/08/attacking-ssl-vpn-part-2-breaking-the-fortigate-ssl-vpn.html

3.3. Правила защиты и стандарты для пользователей

Если по причине недостаточных знаний о степени риска или по принципу предосторожности государство запрещает устаревшую технологию (своего рода «плановое устаревание»), чтобы предотвратить использование старого оборудования, которое можно взломать (Zech, 2021), то оно перекладывает нагрузку на своих граждан, вынуждая их заменять устройства, чтобы избежать штрафов и ответственности. Таким образом, например, компания, сдающая в аренду парк беспилотных автомобилей, может обанкротиться, если не сможет вложить деньги в закупку новых машин. Это может произойти как с крупной транспортной компанией, владеющей беспилотными грузовиками, так и с отдельной семьей, пользующейся одним автомобилем.

Если, напротив, никаких запретов не существует, но бремя исполнения законов также лежит на пользователях, то большинство из них не будет иметь достаточной защиты. Фактически, даже в случае системы обязательного страхования, весьма вероятно, что большое количество людей не получит адекватной компенсации, так как одно лишь страховое покрытие не достаточно эффективно (Samu, 2014).

С точки зрения кибербезопасности, было бы опасно возлагать на пользователей ответственность за использование необновленного устройства в основном потому, что во многих ситуациях принцип ответственности неэффективен. Известно, что степень законопослушности меняется в зависимости от множества обстоятельств, среди которых личные, культурные, социальные факторы и даже возраст (Ma, 1985).

В эпоху, когда, вероятно, многие люди будут пользоваться роботами (например, беспилотными автомобилями), слишком высокими становятся риски от использования устаревших и необновленных устройств, которые могут продаваться на черном рынке тем, кто не может себе позволить новый продукт; это особенно касается ситуаций, когда эти устройства могут приводить к повреждениям и смерти. Беспилотные грузовики могут угонять и использовать для террористических атак (Gallese, 2018), угрожая множеству людей. Нежелательно дозволять пользователям решать, утилизировать старый грузовик или нет. Более того, даже если ЕС введет такую систему, в любой другой стране останутся риски кибератак.

Кроме того, даже если рядовой пользователь готов подчиняться закону, он не способен диагностировать поломку или дефект и исправить его. В самом обычном случае люди, незнакомые с технологиями, будут наиболее уязвимы для атак, даже не осознавая этого. С другой стороны, продвинутые пользователи могут собрать робота в домашних условиях, используя различное программное обеспечение, даже с открытым исходным кодом, которое было опубликовано с другой целью. В этом случае программист не несет ответственности за результат постройки робота, который функционирует через некую комбинацию различного программного обеспечения; ответственным будет пользователь. Разумеется, тот же принцип должен применяться при использовании роботов, которые запрещены законом, сняты с производства или не имеют лицензии.

3.4. Сбалансированная система защищенности и стандартов

Нагрузка, налагаемая нормами безопасности и защищенности, должна распределяться на законодательном уровне между производителями, пользователями и государством, поскольку некоторые меры являются очень дорогостоящими или трудновыполнимыми для пользователей или производителей, тогда как другие для них доступны.

Как только постоянные проверки безопасности на основе анализа степени риска выявили угрозу, министерство может выпустить различные регламенты и принудить потребителей, производителей и даже государственные органы исполнять их, как этого требует ситуация. Например, если на придорожный рекламный щит помещен вредоносный цветовой патч, представляющий непосредственную угрозу для безопасности на дороге, то у производителя не будет достаточно времени, чтобы создать обновление программного обеспечения (если это вообще возможно), поэтому мэр будет обязан перекрыть движение на дороге и в течение нескольких часов удалить объект.

С другой стороны, если разработка обновлений станет обязательной для производителя, можно предусмотреть также обязательную, где это возможно, замену устаревшего программного обеспечения пользователями или регулярное прохождение технического осмотра роботов. Обязанности производителя могут устанавливаться по остаточному принципу, т. е. в тех случаях, которые не могут контролироваться ни государством, ни пользователем.

Это решение не сильно отличается от функционирования других регламентов безопасности: когда новый закон или постановление о чрезвычайной ситуации вводится, чтобы избежать угрозы здоровью, обязанность по его исполнению иногда ложится на пользователя (например, в Италии обязательное устройство, контролирующее посадку ребенка в автомобильном кресле), иногда на производителя (ст. 19 Регламента (Еврокомиссии) № 178/2002 об удалении продуктов питания с рынка). Бремя расходов на исполнение регламента может быть также смягчено через государственное финансирование и другие льготы для производителей и пользователей.

Особое внимание следует уделить в случае возможной угрозы демократии, например, обеспечения национальной безопасности (Allen & Chan, 2017) или выборов. Используемые для этих целей устройства должны регулироваться отдельной и более строгой отраслью права, а контроль над ними должен быть обязанностью государства. Мы считаем, что ни один инструмент, каким-либо образом используемый в демократическом процессе, не должен оставаться в руках частных компаний[26], особенно в эпоху, когда демократия и так находится под угрозой из-за злоупотреблений в области персональных (больших) данных, социальных сетей и фейковых новостей (Manheim & Kaplan, 2019; Persily, 2017; Heawood, 2018; Helbing et al., 2018).

Правила безопасности и защищенности должны применяться также для того, чтобы не позволить производителям выпустить на рынок продукты, обучающиеся в течение всего жизненного цикла, которые могут быть опасными, либо установить особые условия выпуска (обязательное страхование или лицензирование, регистрация и т. д.).

Абсолютно приемлемым является мнение о постоянном присутствии надзора со стороны человека, однако следует учитывать, что с усилением интегрированности технологий в повседневную жизнь люди все больше полагаются на них, не задумываясь о таком надзоре и фактически не осуществляя его; это явление называется ошибкой автоматизации.

---

[26] Это включает возможность распространять политические материалы и рекламу через социальные сети, приложения для смартфонов и домашние устройства, работающие на алгоритмах искусственного интеллекта, и использовать профили пользователей для влияния на их поведение. Скандал с компанией Cambridge Analytica показал насущную необходимость выработать нормативное решение для моделей ИИ, способных повлиять на демократические институты.

4. Принципы закона о киберустойчивости

Новый подход, сформулированный в проекте закона о киберустойчивости, отличается ограниченностью, поскольку он не специфичен для систем ИИ или роботов, что исключает из сферы его действия медицинские устройства, беспилотные автомобили и дроны. Однако он представляет собой важный шаг в направлении развития безопасного и защищенного рынка устройств с цифровыми элементами.

Основные требования, перечисленные в приложениях, постулируют, что такие устройства должны соответствовать ряду правил, а именно:

– включать безопасную по умолчанию конфигурацию и возможность промышленной перезагрузки;

– включать меры защиты устройства от несанкционированного доступа;

– распространять меры обеспечения конфиденциальности и целостности данных на неперсональные данные;

– распространять принципы адекватности, ограничения назначения и минимизации данных на неперсональные данные;

– применять меры киберустойчивости для сохранения существенных функций устройств;

– минимизировать влияние на доступность услуг, предоставляемых другими устройствами или сетями;

– ограничивать поверхность атаки;

– применять соответствующие сдерживающие механизмы и технологии эксплуатации для снижения влияния аварий;

– вести журналы работы;

– обеспечивать обновление систем безопасности на пять лет.

Как и закон об искусственном интеллекте, новый проект также представляет подробный список требований, таких как ведение технической документации, информирование пользователей и оценка рисков.

Предусмотрено особое ограничение для бета-продуктов: «Государства-члены не препятствуют доступу к незавершенному программному обеспечению, которое не соответствует данному Регламенту, если доступ к указанному программному обеспечению предоставляется на ограниченный период времени с целью тестирования и данное программное обеспечение несет на себе четкое указание о несоответствии данному Регламенту и предоставлении исключительно с целью тестирования». Это положение может быть распространено на устройства, разрабатываемые в исследовательских целях, хотя это не указано явным образом в основном тексте проекта.

В пункте 10 поясняется: «Для защиты инноваций или исследований данный Регламент не распространяется на бесплатное программное обеспечение и программное обеспечение с открытым исходным кодом, разрабатываемое или поставляемое вне рамок коммерческой деятельности. Это относится, в частности, к программному обеспечению, включая исходный код и модифицированные версии, которое открыто распространяется, используется, модифицируется и перераспределяется. В контексте программного обеспечения коммерческая деятельность характеризуется не только установлением цены на продукт, но также установлением цены на техническое обслуживание, предоставление программной платформы, через которую производитель монетизирует иные услуги, или использованием персональных

данных по причинам, отличным от повышения безопасности, совместимости или взаимозаменяемости программного обеспечения».

При этом термин «делать доступным на рынке» относится также к устройствам, предоставляемым бесплатно: «[делать доступным на рынке] означает любые поставки продукта с цифровыми элементами для распространения или использования на рынок Евросоюза в процессе коммерческой деятельности, будь то за плату или бесплатно», поэтому ученые, работающие в коллаборации с промышленностью, должны проявлять осторожность.

Выводы

Выработка полного и согласованного законодательства по робототехнике в Европе еще далека от завершения, учитывая быстрое развитие ИИ и многообразие типов устройств, которые могут использоваться в самых разных сферах нашей жизни. Каждая такая сфера имеет свои особенности, поэтому желательно наряду с общими законами – об искусственном интеллекте и о киберустойчивости – задействовать множество других правовых инструментов, мер, инструкций, стандартов, относящихся к различным отраслям.

Первый шаг в разработке всеобъемлющей сбалансированной отрасли права, регулирующей ИИ через закон о киберустойчивости и его нормы безопасности и защищенности, – это установление минимальных технических стандартов, которым должно соответствовать устройство с искусственным интеллектом, прежде чем попадет на рынок Евросоюза. Предотвращение кибератак в важнейших областях и защита фундаментальных прав граждан должны быть главной заботой законодателей. Закон об ИИ и закон о киберустойчивости являются первым шагом в этом направлении, однако их общие принципы не достаточно детализированы, чтобы служить руководством для программистов в их практическом применении.

После того как законодатели разрешат применять все подходящие стандарты и технические регламенты, чтобы гарантировать, что на рынок попадают только безопасные и защищенные продукты, они должны тщательно оценить, в каких случаях на рынок могут быть допущены модели, обучающиеся в течение всего жизненного цикла, и, где это возможно, оставить технологический прогресс прерогативой научных исследований и государственных институтов под центральным наблюдением стран – членов ЕС; особенно это касается опасных продуктов.

В сфере роботов на основе машинного обучения на общеевропейском уровне необходимо специальное техническое регулирование, поскольку существующие нормы неадекватны потребности в регулировании всех вопросов, которые могут возникнуть после выпуска таких устройств на рынок. В частности, учитывая разнообразие подобных технологий, в каждой отрасли должны применяться отдельные нормы. Необходимо также учитывать убытки и риски, возникающие вследствие задержки с принятием новых регламентов в ответ на потребности рынка.

Список литературы

Aladin, D., Varlamov, O., Chuvikov, D., Chernenkiy, V., Smelkova, E., & Baldin, A. (2019). Logic-based artificial intelligence in systems for monitoring the enforcing traffic regulations. In *IOP Conference Series: Materials Science and Engineering* (Vol. 534, p. 012025). IOP Publishing. https://doi.org/10.1088/1757-899x/534/1/012025

Allen, G., & Chan, T. (2017). *Artificial intelligence and national security*. Belfer Center for Science and International Affairs Cambridge, MA.

Ames, H., Mingolla, E., Sohail, A., Chandler, B., Gorchetchnikov, A., L´eveill´e, J., Livitz, G., & Versace, M. (2012). The animat: New frontiers in whole brain modeling. *IEEE pulse*, *3*(1), 47–50. https://doi.org/10.1109/mpul.2011.2175638

Amidei, A. (2017). Robotica intelligente e responsabilit`a: profili e prospettive evolutive del quadro normativo europeo. In U. Ruffolo (ed.), *Intelligenza Artificiale e responsabilit`a, Responsabilit`a Comunicazione Impresa* (Vol. 20, Giuffr`e Editore, pp. 63–106).

Barreno, M., Nelson, B., Joseph, A. D., & Tygar, J. D. (2010). The security of machine learning. *Machine Learning, 81*(2), 121–148. https://doi.org/10.1007/s10994-010-5188-5

Bekey, G. A. (2012). Current trends in robotics: Technology and ethics. In *Robot ethics: the ethical and social implications of robotics* (pp. 17–34). The MIT Press, Cambridge.

Bertolini, A. (2013). Robots as products: the case for a realistic analysis of robotic applications and liability rules. *Law, Innovation and Technology*, *5*(2), 214–247. https://doi.org/10.5235/17579961.5.2.214

Bi, S. (2020). Intelligent system for English translation using automated knowledge base. *Journal of Intelligent & Fuzzy Systems*, *39*(4), 5057–5066. https://doi.org/10.3233/jifs-179991

Botros, N. M., & Abdul-Aziz, M. (1994). Hardware implementation of an artificial neural network using field programmable gate arrays (FPGA's). *IEEE Transactions on Industrial Electronics*, *41*(6), 665–667. https://doi.org/10.1109/41.334585

Buchanan, B. (2016). The life cycles of cyber threats. *Survival*, *58*(1), 39–58. https://doi.org/10.1080/00396338.2016.1142093

Chakkravarthy, R. (2019). Artificial intelligence for construction safety. *Professional Safety*, *64*(1), 46.

Chakraborti, T., Kulkarni, A., Sreedharan, S., Smith, D. E., & Kambhampati, S. (2021). Explicability? Legibility? Predictability? Transparency? Privacy? Security? The Emerging Landscape of Interpretable Agent Behavior. *Proceedings of the International Conference on Automated Planning and Scheduling*, *29*, 86–96. https://doi.org/10.1609/icaps.v29i1.3463

Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications*, *115*(9), 31–41. https://doi.org/10.5120/20182-2402

De Swarte, T., Boufous, O., & Escalle, P. (2019). Artificial intelligence, ethics and human values: the cases of military drones and companion robots. *Artificial Life and Robotics*, *24*(3), 291–296. https://doi.org/10.1007/s10015-019-00525-1

Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and Human-Robot Co-working. *Procedia Computer Science*, *158*, 688–695. https://doi.org/10.1016/j.procs.2019.09.104

Dietterich, T. G., & Kong, E. B. (1995). *Machine learning bias, statistical bias, and statistical variance of decision tree algorithms*. Technical report, Department of Computer Science, Oregon State University.

Dijkstra, E. W. (1982). *Selected writings on computing: a personal perspective*. Springer Science & Business Media.

Ebers, M., & Navas, S. (2020). *Algorithms and Law*. Cambridge University Press.

European Parliament. (2017). *European Parliament resolution of 16 February 2017 with recommendations to the commission on civil law rules on robotics* (2015/2103(INL)).

Falletti, E. (2020). Decisioni automatizzate e diritto alla spiegazione: alcune riflessioni comparatistiche. *Il diritto dell'informazione e dell'informatica*, *3*, 169–206.

Funkhouser, K. (2013). Paving the road ahead: autonomous vehicles, products liability, and the need for a new approach. *Utah Law Review*, 437.

Gallese, C. (2018). Prospettive di riforma del diritto internazionale privato giapponese. In M. Cestari, G. Coci, D. Moro, A. Specchio (Eds.), *Orizzonti giapponesi: ricerche, idee, prospettive* (pp. 185–186). Aracne editrice, Roma.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining Explanations: An Overview of Interpretability of Machine Learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 80–89). https://doi.org/10.1109/dsaa.2018.00018

Greene, D., Hoffmann, A. L., & Stark, L. (2019). Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2019.258

Haselager, W. F. (2005). Robotics, philosophy and the problems of autonomy. *Cognitive Technologies and the Pragmatics of Cognition*, *13*(3), 515–532. https://doi.org/10.1075/pc.13.3.07has

Heawood, J. (2018). Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity*, *23*(4), 429–434. https://doi.org/10.3233/ip-180009

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., van den Hoven, J., Zicari, R. V., & Zwitter, A. (2018). Will Democracy Survive Big Data and Artificial Intelligence? *Towards Digital Enlightenment*, 73–98. https://doi.org/10.1007/978-3-319-90869-4_7

Hildebrandt, M. (2019). Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning. *Theoretical Inquiries in Law*, *20*(1), 83–121. https://doi.org/10.1515/til-2019-0004

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *WIREs Data Mining and Knowledge Discovery*, *9*(4). https://doi.org/10.1002/widm.1312

Ielmini, D., & Ambrogio, S. (2019). Emerging neuromorphic devices. *Nanotechnology*, *31*(9), 092001. https://doi.org/10.1088/1361-6528/ab554b

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Lamberton, C., Brigo, D., & Hoy, D. (2017). Impact of robotics, RPA and AI on the insurance industry: challenges and opportunities. *Journal of Financial Perspectives*, *4*(1).

Leenes, R., Palmerini, E., Koops, B. J., Bertolini, A., Salvini, P., & Lucivero, F. (2017). Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues. *Law, Innovation and Technology*, *9*(1), 1–44. https://doi.org/10.1080/17579961.2017.1304921

Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2017). Brain Intelligence: Go beyond Artificial Intelligence. *Mobile Networks and Applications*, *23*(2), 368–375. https://doi.org/10.1007/s11036-017-0932-8

Luna, R., Rhine, E., Myhra, M., Sullivan, R., & Kruse, C. S. (2016). Cyber threats to health information systems: A systematic review. *Technology and Health Care*, *24*(1), 1–9. https://doi.org/10.3233/thc-151102

Ma, H. K. (1985). Cross-Cultural Study of the Development of Law-Abiding Orientation. *Psychological Reports*, *57*(3), 967–974. https://doi.org/10.2466/pr0.1985.57.3.967

Manheim, K. M., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy. *Yale Journal of Law and Technology*, *21*, 106–188.

Naur, P., & Randell, B. (Eds.) (1968). *Software Engineering: Report of a conference sponsored by the NATO Science Committee.* Newcastle University.

Nevejans, N. (2016). *European civil law rules in robotics*. Policy Department for Citizens' Rights and Constitutional Affairs.

Nobile, M. S., Vlachou, T., Spolaor, S., Cazzaniga, P., Mauri, G., Pelicci, P. G., & Besozzi, D. (2019). ProCell: Investigating cell proliferation with Swarm Intelligence. *2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (pp. 1–8). https://doi.org/10.1109/cibcb.2019.8791468

O'Sullivan, S., Nevejans, N., Allen, C., Blyth, A., Leonard, S., Pagallo, U., Holzinger, K., Holzinger, A., Sajid, M. I., & Ashrafian, H. (2019). Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, *15*(1), e1968. https://doi.org/10.1002/rcs.1968

Ozkan, B. E., & Bulkan, S. (2019). Hidden Risks to Cyberspace Security from Obsolete COTS Software. *2019 11th International Conference on Cyber Conflict (CyCon)* (pp. 1–19). https://doi.org/10.23919/cycon.2019.8756990

Pan, C., Wang, C. Y., Liang, S. J., Wang, Y., Cao, T., Wang, P., Wang, C., Wang, S., Cheng, B., Gao, A., Liu, E., Watanabe, K., Taniguchi, T., & Miao, F. (2020). Reconfigurable logic and neuromorphic circuits based on electrically tunable two-dimensional homojunctions. *Nature Electronics*, *3*(7), 383–390. https://doi.org/10.1038/s41928-020-0433-9

Pandian, D. A. P. (2019). Artificial intelligence application in smart warehousing environment for automated logistics. *Journal of Artificial Intelligence and Capsule Networks*, *1*(2), 63–72. https://doi.org/10.36548/jaicn.2019.2.002

Papetti, D. M., Spolaor, S., Besozzi, D., Cazzaniga, P., Antoniotti, M., & Nobile, M. S. (2020). On the automatic calibration of fully analogical spiking neuromorphic chips. *2020 International Joint Conference on Neural Networks (IJCNN)*. https://doi.org/10.1109/ijcnn48605.2020.9206654

Persily, N. (2017). Can Democracy Survive the Internet? *Journal of Democracy*, *28*(2), 63–76. https://doi.org/10.1353/jod.2017.0025

Ranjan, A., Janai, J., Geiger, A., & Black, M. (2019). Attacking Optical Flow. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 2404–2413). https://doi.org/10.1109/iccv.2019.00249

Rao, Q., & Frtunikj, J. (2018). Deep learning for self-driving cars. *Proceedings of the 1st International Workshop on Software Engineering for AI in Autonomous Systems* (pp. 35–38). https://doi.org/10.1145/3194085.3194087

Raza, M., Awais, M., Ali, K., Aslam, N., Paranthaman, V. V., Imran, M., & Ali, F. (2020). Establishing effective communications in disaster affected areas and artificial intelligence based detection using social media platform. *Future Generation Computer Systems*, *112*, 1057–1069. https://doi.org/10.1016/j.future.2020.06.040

Richards, N. M., & Smart, W. D. (2016). How should the law think about robots? In R. Calo, M. A. Froomkin, I. Kerr (Eds.). *Robot law*. Edward Elgar.

Roll, I., & Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Romanishin, J. W., Gilpin, K., & Rus, D. (2013). M-blocks: Momentum-driven, magnetic modular robots. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4288–4295). https://doi.org/10.1109/iros.2013.6696971

Rundo, L., Tangherloni, A., Nobile, M. S., Militello, C., Besozzi, D., Mauri, G., & Cazzaniga, P. (2019). MedGA: A novel evolutionary method for image enhancement in medical imaging systems. *Expert Systems With Applications*, *119*, 387–399. https://doi.org/10.1016/j.eswa.2018.11.013

Samu, S. (2014). The effectiveness of compulsory motor insurance in Zimbabwe. *Journal of Strategic Studies: A Journal of the Southern Bureau of Strategic Studies Trust*, *5*(1), 45–60.

Santosuosso, A., Boscarato, C., & Caroleo, F. (2012). Robot e diritto: una prima ricognizione. *La Nuova Giurisprudenza Commentata*, *494*.

Simoens, P., Mahieu, C., Ongenae, F., De Backere, F., De Pestel, S., Nelis, J., De Turck, F., Elprama, S. A., Kilpi, K., Jewell, C., & Jacobs, A. (2016). Internet of Robotic Things: Context-Aware and Personalized Interventions of Assistive Social Robots (Short Paper). *2016 5th IEEE International Conference on Cloud Networking (Cloudnet)* (pp. 204–207). https://doi.org/10.1109/cloudnet.2016.27

Smith, B. Walker (2016). Lawyers and engineers should speak the same robot language. In R. Calo, M. A. Froomkin, I. Kerr (Eds.), *Robot law* (pp. 78–101). Edward Elgar Publishing.

Sourdin, T. (2018). Judge v Robot? Artificial Intelligence and Judicial Decision-Making. *University of New South Wales Law Journal*, *41*(4), 25167–25177. https://doi.org/10.53637/zgux2213

Wang, W., & Siau, K. (2019). Artificial Intelligence, Machine Learning, Automation, Robotics, Future of Work and Future of Humanity. *Journal of Database Management*, *30*(1), 61–79. https://doi.org/10.4018/jdm.2019010104

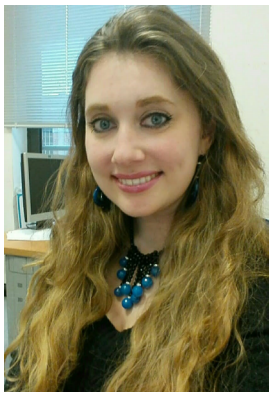
Yano, K. (2017). How artificial intelligence will change HR. *People & Strategy*, *40*(3), 42–47.

Yuniarthe, Y. (2017). Application of Artificial Intelligence (AI) in Search Engine Optimization (SEO). In *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)* (pp. 96–101). https://doi.org/10.1109/icsiit.2017.15

Zech, H. (2021). Liability for AI: public policy considerations. *ERA Forum*, *22*(1), 147–158. https://doi.org/10.1007/s12027-020-00648-0

Zhang, X., Upton, O., Beebe, N. L., & Choo, K. K. R. (2020). IoT Botnet Forensics: A Comprehensive Digital Forensic Case Study on Mirai Botnet Servers. *Forensic Science International: Digital Investigation*, *32*, 300926. https://doi.org/10.1016/j.fsidi.2020.300926

Сведения об авторе



Галлезе-Нобиле Кьяра – доктор наук, научный сотрудник (постдок) по управлению исследовательскими данными, Эйндховенский технологический университет (Эйндховен, Королевство Нидерландов); научный сотрудник (постдок) департамента математики и наук о земле, Университет Триеста (Триест, Итальянская Республика)
Адрес: а/я 513 5600 МБ Эйндховен, Королевство Нидерландов
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57222726276
Web of Science Researcher ID:
https://www.webofscience.com/wos/author/record/AGE-9594-2022
ORCID ID: https://orcid.org/0000-0001-8194-0261
Google Scholar ID: https://scholar.google.com/citations?user=Vmoen8UAAAAJ

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Финансирование

Исследование выполнено при поддержке проекта UNI 4 JUSTICE.

История статьи

Дата поступления – 13 октября 2022 г.
Дата одобрения после рецензирования – 4 ноября 2022 г.
Дата принятия к опубликованию – 6 марта 2023 г.
Дата онлайн-размещения – 10 марта 2023 г.

Research article
DOI: https://doi.org/10.21202/jdtl.2023.2

Regulating Smart Robots and Artificial Intelligence in the European Union

Chiara Gallese Nobile

Eindhoven University of Technology
Eindhoven, Kingdom of the Netherlands;
University of Trieste
Trieste, Italian Republic

Keywords

Artificial intelligence,
cybersecurity,
digital technologies,
European Union,
law,
legislation,
machine learning,
regulation,
robot,
robotics

Abstract

Objective: In recent years, the need for regulation of robots and Artificial Intelligence has become apparent in Europe. European Union needs a standardized regulation that will ensure a high level of security in robotics systems to prevent potential breaches. Therefore a new regulation should make clear that it is the responsibility of producers to identify the blind spots in these systems, exposing their flaws, or, when a vulnerability is discovered in a later stage, to update the system even if that model is not on the market anymore. This article aims at suggesting some possible revisions of the existing legal provisions in the EU.

Methods: The author employed the Kestemont legal methodology, analyzing legal text, comparing them, and connecting them with technical elements regarding smart robots, resulting in the highlighting of the critical provisions to be updated

Results: This article suggests some revisions to the existing regulatory proposals: according to the author, although the AI Act and the Cyber-resilience Act represent a first step towards this direction, their general principles are not sufficiently detailed to guide programmers on how to implement them in practice, and policymakers should carefully assess in what cases lifelong learning models should be allowed to the market. The author suggests that the current proposal regarding mandatory updates should be expanded, as five years are a short time frame that would not cover the risks associated with long-lasting products, such as vehicles.

© Gallese Nobile C., 2023

This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Scientific novelty: The author has examined the existing regulatory framework regarding AI systems and devices with digital elements, highlighted the risks of the current legal framework, and suggested possible amendments to the existing regulatory proposals

Practical significance: The article can be employed to update the existing proposals for the AI Act and the Cyber-resilience Act

For citation

Gallese Nobile, C. (2023). Regulating Smart Robots and Artificial Intelligence in the European Union. *Journal of Digital Technologies and Law*, *1*(1), 33–61. https://doi.org/10.21202/jdtl.2023.2

References

Aladin, D., Varlamov, O., Chuvikov, D., Chernenkiy, V., Smelkova, E., & Baldin, A. (2019). Logic-based artificial intelligence in systems for monitoring the enforcing traffic regulations. In *IOP Conference Series: Materials Science and Engineering* (Vol. 534, p. 012025). IOP Publishing. https://doi.org/10.1088/1757-899x/534/1/012025

Allen, G., & Chan, T. (2017). *Artificial intelligence and national security*. Belfer Center for Science and International Affairs Cambridge, MA.

Ames, H., Mingolla, E., Sohail, A., Chandler, B., Gorchetchnikov, A., L´eveill´e, J., Livitz, G., & Versace, M. (2012). The animat: New frontiers in whole brain modeling. *IEEE pulse*, *3*(1), 47–50. https://doi.org/10.1109/mpul.2011.2175638

Amidei, A. (2017). Robotica intelligente e responsabilit`a: profili e prospettive evolutive del quadro normativo europeo. In U. Ruffolo (ed.), *Intelligenza Artificiale e responsabilit`a, Responsabilit`a Comunicazione Impresa* (Vol. 20, Giuffr`e Editore, pp. 63–106).

Barreno, M., Nelson, B., Joseph, A. D., & Tygar, J. D. (2010). The security of machine learning. *Machine Learning, 81*(2), 121–148. https://doi.org/10.1007/s10994-010-5188-5

Bekey, G. A. (2012). Current trends in robotics: Technology and ethics. In *Robot ethics: the ethical and social implications of robotics* (pp. 17–34). The MIT Press, Cambridge.

Bertolini, A. (2013). Robots as products: the case for a realistic analysis of robotic applications and liability rules. *Law, Innovation and Technology*, *5*(2), 214–247. https://doi.org/10.5235/17579961.5.2.214

Bi, S. (2020). Intelligent system for English translation using automated knowledge base. *Journal of Intelligent & Fuzzy Systems*, *39*(4), 5057–5066. https://doi.org/10.3233/jifs-179991

Botros, N. M., & Abdul-Aziz, M. (1994). Hardware implementation of an artificial neural network using field programmable gate arrays (FPGA's). *IEEE Transactions on Industrial Electronics*, *41*(6), 665–667. https://doi.org/10.1109/41.334585

Buchanan, B. (2016). The life cycles of cyber threats. *Survival*, *58*(1), 39–58. https://doi.org/10.1080/00396338.2016.1142093

Chakkravarthy, R. (2019). Artificial intelligence for construction safety. *Professional Safety*, *64*(1), 46.

Chakraborti, T., Kulkarni, A., Sreedharan, S., Smith, D. E., & Kambhampati, S. (2021). Explicability? Legibility? Predictability? Transparency? Privacy? Security? The Emerging Landscape of Interpretable Agent Behavior. *Proceedings of the International Conference on Automated Planning and Scheduling*, *29*, 86–96. https://doi.org/10.1609/icaps.v29i1.3463

Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications*, *115*(9), 31–41. https://doi.org/10.5120/20182-2402

De Swarte, T., Boufous, O., & Escalle, P. (2019). Artificial intelligence, ethics and human values: the cases of military drones and companion robots. *Artificial Life and Robotics*, *24*(3), 291–296. https://doi.org/10.1007/s10015-019-00525-1

Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and Human-Robot Co-working. *Procedia Computer Science*, *158*, 688–695. https://doi.org/10.1016/j.procs.2019.09.104

Dietterich, T. G., & Kong, E. B. (1995). *Machine learning bias, statistical bias, and statistical variance of decision tree algorithms*. Technical report, Department of Computer Science, Oregon State University.

Dijkstra, E. W. (1982). *Selected writings on computing: a personal perspective*. Springer Science & Business Media.
Ebers, M., & Navas, S. (2020). *Algorithms and Law*. Cambridge University Press.
European Parliament. (2017). *European Parliament resolution of 16 February 2017 with recommendations to the commission on civil law rules on robotics* (2015/2103(INL)).
Falletti, E. (2020). Decisioni automatizzate e diritto alla spiegazione: alcune riflessioni comparatistiche. *Il diritto dell'informazione e dell'informatica*, *3*, 169–206.
Funkhouser, K. (2013). Paving the road ahead: autonomous vehicles, products liability, and the need for a new approach. *Utah Law Review*, 437.
Gallese, C. (2018). Prospettive di riforma del diritto internazionale privato giapponese. In M. Cestari, G. Coci, D. Moro, A. Specchio (Eds.), *Orizzonti giapponesi: ricerche, idee, prospettive* (pp. 185–186). Aracne editrice, Roma.
Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining Explanations: An Overview of Interpretability of Machine Learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 80–89). https://doi.org/10.1109/dsaa.2018.00018
Greene, D., Hoffmann, A. L., & Stark, L. (2019). Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2019.258
Haselager, W. F. (2005). Robotics, philosophy and the problems of autonomy. *Cognitive Technologies and the Pragmatics of Cognition*, *13*(3), 515–532. https://doi.org/10.1075/pc.13.3.07has
Heawood, J. (2018). Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity*, *23*(4), 429–434. https://doi.org/10.3233/ip-180009
Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., van den Hoven, J., Zicari, R. V., & Zwitter, A. (2018). Will Democracy Survive Big Data and Artificial Intelligence? *Towards Digital Enlightenment*, 73–98. https://doi.org/10.1007/978-3-319-90869-4_7
Hildebrandt, M. (2019). Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning. *Theoretical Inquiries in Law*, *20*(1), 83–121. https://doi.org/10.1515/til-2019-0004
Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *WIREs Data Mining and Knowledge Discovery*, *9*(4). https://doi.org/10.1002/widm.1312
Ielmini, D., & Ambrogio, S. (2019). Emerging neuromorphic devices. *Nanotechnology*, *31*(9), 092001. https://doi.org/10.1088/1361-6528/ab554b
Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004
Lamberton, C., Brigo, D., & Hoy, D. (2017). Impact of robotics, RPA and AI on the insurance industry: challenges and opportunities. *Journal of Financial Perspectives*, *4*(1).
Leenes, R., Palmerini, E., Koops, B. J., Bertolini, A., Salvini, P., & Lucivero, F. (2017). Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues. *Law, Innovation and Technology*, *9*(1), 1–44. https://doi.org/10.1080/17579961.2017.1304921
Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2017). Brain Intelligence: Go beyond Artificial Intelligence. *Mobile Networks and Applications*, *23*(2), 368–375. https://doi.org/10.1007/s11036-017-0932-8
Luna, R., Rhine, E., Myhra, M., Sullivan, R., & Kruse, C. S. (2016). Cyber threats to health information systems: A systematic review. *Technology and Health Care*, *24*(1), 1–9. https://doi.org/10.3233/thc-151102
Ma, H. K. (1985). Cross-Cultural Study of the Development of Law-Abiding Orientation. *Psychological Reports*, *57*(3), 967–974. https://doi.org/10.2466/pr0.1985.57.3.967
Manheim, K. M., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy. *Yale Journal of Law and Technology*, *21*, 106–188.
Naur, P., & Randell, B. (Eds.) (1968). *Software Engineering: Report of a conference sponsored by the NATO Science Committee.* Newcastle University.
Nevejans, N. (2016). *European civil law rules in robotics*. Policy Department for Citizens' Rights and Constitutional Affairs.
Nobile, M. S., Vlachou, T., Spolaor, S., Cazzaniga, P., Mauri, G., Pelicci, P. G., & Besozzi, D. (2019). ProCell: Investigating cell proliferation with Swarm Intelligence. *2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (pp. 1–8). https://doi.org/10.1109/cibcb.2019.8791468
O'Sullivan, S., Nevejans, N., Allen, C., Blyth, A., Leonard, S., Pagallo, U., Holzinger, K., Holzinger, A., Sajid, M. I., & Ashrafian, H. (2019). Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, *15*(1), e1968. https://doi.org/10.1002/rcs.1968

Ozkan, B. E., & Bulkan, S. (2019). Hidden Risks to Cyberspace Security from Obsolete COTS Software. *2019 11th International Conference on Cyber Conflict (CyCon)* (pp. 1–19). https://doi.org/10.23919/cycon.2019.8756990

Pan, C., Wang, C. Y., Liang, S. J., Wang, Y., Cao, T., Wang, P., Wang, C., Wang, S., Cheng, B., Gao, A., Liu, E., Watanabe, K., Taniguchi, T., & Miao, F. (2020). Reconfigurable logic and neuromorphic circuits based on electrically tunable two-dimensional homojunctions. *Nature Electronics*, *3*(7), 383–390. https://doi.org/10.1038/s41928-020-0433-9

Pandian, D. A. P. (2019). Artificial intelligence application in smart warehousing environment for automated logistics. *Journal of Artificial Intelligence and Capsule Networks*, *1*(2), 63–72. https://doi.org/10.36548/jaicn.2019.2.002

Papetti, D. M., Spolaor, S., Besozzi, D., Cazzaniga, P., Antoniotti, M., & Nobile, M. S. (2020). On the automatic calibration of fully analogical spiking neuromorphic chips. *2020 International Joint Conference on Neural Networks (IJCNN)*. https://doi.org/10.1109/ijcnn48605.2020.9206654

Persily, N. (2017). Can Democracy Survive the Internet? *Journal of Democracy*, *28*(2), 63–76. https://doi.org/10.1353/jod.2017.0025

Ranjan, A., Janai, J., Geiger, A., & Black, M. (2019). Attacking Optical Flow. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 2404–2413). https://doi.org/10.1109/iccv.2019.00249

Rao, Q., & Frtunikj, J. (2018). Deep learning for self-driving cars. *Proceedings of the 1st International Workshop on Software Engineering for AI in Autonomous Systems* (pp. 35–38). https://doi.org/10.1145/3194085.3194087

Raza, M., Awais, M., Ali, K., Aslam, N., Paranthaman, V. V., Imran, M., & Ali, F. (2020). Establishing effective communications in disaster affected areas and artificial intelligence based detection using social media platform. *Future Generation Computer Systems*, *112*, 1057–1069. https://doi.org/10.1016/j.future.2020.06.040

Richards, N. M., & Smart, W. D. (2016). How should the law think about robots? In R. Calo, M. A. Froomkin, I. Kerr (Eds.). *Robot law*. Edward Elgar.

Roll, I., & Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Romanishin, J. W., Gilpin, K., & Rus, D. (2013). M-blocks: Momentum-driven, magnetic modular robots. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4288–4295). https://doi.org/10.1109/iros.2013.6696971

Rundo, L., Tangherloni, A., Nobile, M. S., Militello, C., Besozzi, D., Mauri, G., & Cazzaniga, P. (2019). MedGA: A novel evolutionary method for image enhancement in medical imaging systems. *Expert Systems With Applications*, *119*, 387–399. https://doi.org/10.1016/j.eswa.2018.11.013

Samu, S. (2014). The effectiveness of compulsory motor insurance in Zimbabwe. *Journal of Strategic Studies: A Journal of the Southern Bureau of Strategic Studies Trust*, *5*(1), 45–60.

Santosuosso, A., Boscarato, C., & Caroleo, F. (2012). Robot e diritto: una prima ricognizione. *La Nuova Giurisprudenza Commentata*, *494*.

Simoens, P., Mahieu, C., Ongenae, F., De Backere, F., De Pestel, S., Nelis, J., De Turck, F., Elprama, S. A., Kilpi, K., Jewell, C., & Jacobs, A. (2016). Internet of Robotic Things: Context-Aware and Personalized Interventions of Assistive Social Robots (Short Paper). *2016 5th IEEE International Conference on Cloud Networking (Cloudnet)* (pp. 204–207). https://doi.org/10.1109/cloudnet.2016.27

Smith, B. Walker (2016). Lawyers and engineers should speak the same robot language. In R. Calo, M. A. Froomkin, I. Kerr (Eds.), *Robot law* (pp. 78–101). Edward Elgar Publishing.

Sourdin, T. (2018). Judge v Robot? Artificial Intelligence and Judicial Decision-Making. *University of New South Wales Law Journal*, *41*(4), 25167–25177. https://doi.org/10.53637/zgux2213

Wang, W., & Siau, K. (2019). Artificial Intelligence, Machine Learning, Automation, Robotics, Future of Work and Future of Humanity. *Journal of Database Management*, *30*(1), 61–79. https://doi.org/10.4018/jdm.2019010104

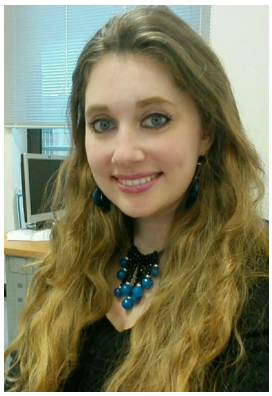
Yano, K. (2017). How artificial intelligence will change HR. *People & Strategy*, *40*(3), 42–47.

Yuniarthe, Y. (2017). Application of Artificial Intelligence (AI) in Search Engine Optimization (SEO). In *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)* (pp. 96–101). https://doi.org/10.1109/icsiit.2017.15

Zech, H. (2021). Liability for AI: public policy considerations. *ERA Forum*, *22*(1), 147–158. https://doi.org/10.1007/s12027-020-00648-0

Zhang, X., Upton, O., Beebe, N. L., & Choo, K. K. R. (2020). IoT Botnet Forensics: A Comprehensive Digital Forensic Case Study on Mirai Botnet Servers. *Forensic Science International: Digital Investigation*, *32*, 300926. https://doi.org/10.1016/j.fsidi.2020.300926

Author information



Chiara Gallese Nobile – PhD, Researcher (postdoc) of research data management, Eindhoven University of Technology (Eindhoven, the Netherlands), Researcher (postdoc) of the Department of Mathematics and Geosciences at the University of Trieste (Trieste, Italy).
Address: P/O 513 5600 MB Eindhoven, the Netherlands
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57222726276
Web of Science Researcher ID:
https://www.webofscience.com/wos/author/record/AGE-9594-2022
ORCID ID: https://orcid.org/0000-0001-8194-0261
Google Scholar ID: https://scholar.google.com/citations?user=Vmoen8UAAAAJ

Conflict of interest

The author declares no conflict of interest.

Financial disclosure

The research has been funded by the UNI 4 JUSTICE project.

Article history

Date of receipt – October 13, 2022
Date of approval – November 4, 2022
Date of acceptance – March 6, 2023
Date of online placement – March 10, 2023