

## НАУКА, ТЕХНИКА, ОБЩЕСТВО

*А. Грунвальд, Д.В. Ефременко*

### **Цифровая трансформация и социальная оценка техники**

*Грунвальд Армин* – доктор философии, профессор, директор Института социальной оценки техники и системного анализа. Технологический институт Карлсруэ (ФРГ). Руководитель Бюро социальной оценки техники при Бундестаге ФРГ. Федеративная Республика Германия, D-76021, г. Карлсруэ, P.O. Box 3640; e-mail: armin.grunwald@kit.edu

*Ефременко Дмитрий Валерьевич* – доктор политических наук, заместитель директора. Институт научной информации по общественным наукам РАН. Российская Федерация, 117997, г. Москва, Нахимовский проспект, д. 51/21; e-mail: efdv2015@mail.ru

В статье рассматриваются возможности применения социальной оценки техники (Technology Assessment) как особой отрасли комплексных междисциплинарных исследований долгосрочных последствий развития новейших технологий в контексте цифровой трансформации современного общества. Социальная оценка цифровых технологий позволяет осуществлять поддержку процесса принятия политических решений, а также находить варианты урегулирования конфликтов социальных акторов, вызванных процессами цифровизации. При этом необходимо принимать во внимание институциональные эффекты развития цифровых технологий. Принципиальное значение имеет интеграция в процесс социального контроля цифровой трансформации принципов этики и социальной ответственности. Особое внимание в статье уделяется таким направлениям цифровой трансформации как развитие индустрии 4.0, широкое внедрение беспилотных автомобилей и роботов для ухода за тяжелобольными и престарелыми.

**Ключевые слова:** цифровая трансформация, социальная оценка техники, индустрия 4.0, беспилотный автотранспорт, роботизированные системы, ответственность в исследованиях и инновациях

Развитие информационно-коммуникационных технологий на протяжении XX в. и новый качественный переход – цифровая трансформация – не просто затрагивают все сферы человеческой жизни, но и подводят цивилизацию к порогу фундаментальных преобразований самой жизни. Это означает возникновение

комплекса проблем, связанных с формированием и управлением цифровым обществом, устойчивой интеграцией в процесс принимаемых решений ценностей, аспектов морали и социальной ответственности. Речь идет и о новом уровне социальных воздействий научно-технических инноваций, поскольку цифровой переход предполагает управление уже не отдельными объектами (как в случае информатизации), но экосистемами объектов. Само проектирование и управление в условиях цифровизации предполагает не только наличие цифрового следа соответствующего объекта, но и создание его цифрового двойника. Дальнейшие процессы НИОКР будут происходить во все более оцифрованной среде. Из этого также следует, что в эпоху цифровизации значение ценностей в проектировании технологий будущего еще более возрастет в связи с необходимостью формирования все более устойчивых и глубоких взаимодействий между людьми и технологиями. Проектирование человеко-машинных интерфейсов и на индивидуальном уровне, и на уровне социума – не только функциональная проблема, оно затрагивает этические вопросы, включая распределение ответственности, и антропологические проблемы, связанные с самосознанием и самооценкой человека.

### **Новые перспективы социальной оценки техники**

Проблема экспертно-аналитической поддержки научно-технической деятельности в условиях цифрового общества и обеспечения учета интересов и ценностей в данном процессе различных общественных групп обретает в настоящее время особую актуальность. Одним из возможных решений является социальная оценка техники (Technology Assessment / ТА), для обозначения которой в русскоязычной научной литературе также достаточно широко используется термин «оценка техники» [Ефременко, 2002; Горохов, 2015].

Социальная оценка техники – особая отрасль междисциплинарных исследований существующих или потенциальных последствий развития науки и техники, особенностью которой является использование широкого диапазона исследовательских методов и подходов. В качестве инструмента научной поддержки принятия политических решений ТА призвана содействовать рационализации и рефлексивности научно-технической политики, аккумулируя имеющиеся знания о побочных эффектах развития той или иной технологии и способствуя разработке стратегий действий, учитывающих социальные ценности и возможные конфликты интересов.

За полвека своего институционального развития в странах Запада социальная оценка техники проделала большой путь в плане теоретико-методологических подходов, а также модальностей взаимодействия со структурами власти и гражданским обществом. На начальном этапе идеология ТА следовала установкам технологического детерминизма, а основная задача социальной оценки техники сводилась преимущественно к раннему предупреждению и выявлению спектра вероятных негативных последствий технического развития. Считалось, что на уровне принятия управленческих решений это позволит действовать на опережение, блокировать негативные эффекты для общества и природы либо их компенсировать. Стоит отметить, что определенное

возрождение элементов технологического детерминизма происходит и в настоящее время в связи с процессами цифровизации, масштабы которых сопоставимы с природными феноменами, такими как цунами или землетрясения, а скорость не позволяет оказать на них существенное корректирующее воздействие. Более того, без упреждающей подготовки еще на уровне среднего образования к комплексным изменениям, связанным с цифровизацией, последствия могут оказаться разрушительными для многих предприятий и целых отраслей экономики, а также рынка труда [Grunwald, 2019, p. 159]. Ситуация усугубляется тем, что программное обеспечение (например, алгоритмы поисковых систем) начинает играть роль специфического социального института, оказывающего регулятивное воздействие на поведение и предпочтения индивидов и групп, не будучи при этом в достаточной степени транспарентным или демократически легитимным [Brown, Marsden, 2013].

Парадигмальные изменения в подходах социальной оценки техники стали происходить в 1980-е гг. в связи распространением идей социального конструктивизма [Bijker, Hughes, Pinch, 1987]. В некотором смысле можно говорить о росте амбиций ТА, конструктивистская версия которой (Constructive Technology Assessment) означала переориентацию на сопровождение процесса создания новых технологий с точки зрения интеграции в этот процесс социальных потребностей и ценностей [Rip, Misa, Schot, 1995]. Фактически ТА должна начинать оказывать воздействие уже на стадии проектирования новой технологии. Для достижения этой цели потребовалось предложить комбинацию методов, включающую исследовательские подходы, а также интерактивные и коммуникативные методы, позволяющие обеспечить оптимальное взаимодействие заинтересованных сторон, предотвращение конфликтов или их разрешение. Исследовательские методы включают сбор данных, формирование достаточной эмпирической базы для прогнозирования прямых и побочных эффектов, количественную оценку рисков, выявление социально-экономических последствий, учет социальных ценностей, оценку влияния на экологическое равновесие. Интерактивные методы ориентированы на организацию и модерацию диалога заинтересованных сторон, на то, чтобы принять во внимание существующие интересы и опыт стейкхолдеров, насколько возможно обеспечив их участие в процессе принятия управленческих решений.

При таком понимании методов и задач ТА цифровизация, очевидно, становится серьезнейшим вызовом для социальной оценки техники. Нет сомнений, что направляющие процесс цифровой трансформации политические решения, а также программные алгоритмы, заложенные в цифровых технологиях, служат отражением определенных интересов и ценностей. Помимо принципиальных вопросов о легитимности и репрезентативности этих интересов и ценностей, общая проблема состоит в нарастающей сложности последующего социального контроля, в том, что цифровые технологии в их натиске и тотальности превращаются в квазиприродную силу, к которой индивиды и сообщества вынуждены приспосабливаться. Иначе говоря, становление цифрового общества сопровождается радикализацией дилеммы контроля [Collingridge, 1980], согласно которой предсказания эффектов внедрения новой

технологии крайне ненадежны до тех пор, пока это внедрение не состоялось, а когда технология широко внедрена и массово используется, внесение изменений сопряжено с непомерными издержками. К тому же, как показал Г. Коллингридж, чем более значимой по масштабу потенциального воздействия является технология, тем более централизованным оказывается процесс принятия связанных с ней решений, а участвующие в их подготовке эксперты чаще всего предпочитают укрепить свои позиции, оказав проекту поддержку. В то же время основные пользователи или те, на кого технология может оказать прямое или опосредованное воздействие, в большинстве случаев не имеют отношения к принятию соответствующих решений [Collingridge, 1990].

### Индустрия 4.0

Казалось бы, такое положение дел в сфере цифровизации подрывает эффективность усилий специалистов по социальной оценке техники. Однако это не так. Рассмотрим круг проблем, связанных с развитием Индустрии 4.0. Этот термин, равно как и термин «Четвертая промышленная революция», впервые был использован в докладе для правительства ФРГ (2011), посвященном смене парадигмы промышленного производства, обусловленной внедрением киберфизических систем и в особенности Интернета вещей [Kagermann, Lukas, Wahlster, 2011]. Спустя несколько лет с легкой руки председателя Всемирного экономического форума в Давосе К. Шваба идея Четвертой промышленной революции получила глобальную известность [Шваб, 2017]. В русскоязычном научном пространстве для характеристики части этих трансформаций также используется понятие «Шестой технологический уклад», хотя о его полной идентичности «Четвертой промышленной революции» говорить едва ли уместно, поскольку в России речь идет о варианте развития теории экономических циклов Н.Д. Кондратьева, а основной упор в концепции Шестого технологического уклада делается на широком внедрении нанотехнологий [Глазьев, Харитонов, 2009].

Цифровизация промышленного производства в контексте Индустрии 4.0 будет разворачиваться как самоорганизующийся процесс. Объединение технических устройств и инфраструктур в единую сеть в режиме реального времени существенно изменит производственные процессы, бизнес-модели, продукты и услуги, а также характер трудовых отношений. В организационном отношении переход к Индустрии 4.0 должен происходить в соответствии со следующими основными принципами:

1. Сетевое структурирование: люди, машины, технические и сенсорные устройства взаимодействуют и обмениваются данными друг с другом через Интернет вещей и традиционный Интернет.

2. Информационная транспарентность: данные сенсорных устройств обеспечивают постоянное расширение возможностей эффективного функционирования информационных систем цифровых фабрик для создания виртуального образа реального мира («цифровых двойников» реальных объектов), позволяющего выйти на новое качество всех этапов проектирования и производства (переход к т.н. «умным фабрикам» / smart factories).

3. Техническое обеспечение: системы поддержки искусственного интеллекта (ИИ) помогают людям принимать обоснованные решения и быстрее реагировать на возникающие проблемы.

4. Децентрализация решений: производственные системы принимают независимые решения и выполняют свои задачи максимально автономно. Только в исключительных случаях, например, в ситуации сбоя или противоречия целевых установок, они делегируют задачи на верхний уровень, например, человеку, осуществляющему общее руководство соответствующими процессами [Hermann, Pentek, Otto, 2016].

Цель перехода к Индустрии 4.0 состоит в том, чтобы сделать возможным индивидуализированное производство в соответствии с требованиями заказчиков, создав тем самым радикальную альтернативу фордистскому массовому производству. Необходимая в этом случае технология автоматизации в основном должна базироваться на искусственном интеллекте, а ее функциональными особенностями будут самооптимизация, самонастройка, самодиагностика и автономное познание. Помимо создания индивидуализированных продуктов, сокращение затрат времени и повышение эффективности производства должны достигаться за счет обучения в процессе производства и аккумуляции знаний и навыков. Географически распределенные производственные мощности и соответствующие системы планирования и управления должны взаимодействовать автономно и, таким образом, лучше использовать существующие ресурсы. Реализуя эти ожидания, ИИ будет играть решающую роль в обработке данных, процессах обучения и самоорганизации.

Проектирование технологий в данном случае сталкивается со значительными трудностями. Прежде всего, должны быть разработаны технические стандарты и регламенты, обеспечивающие взаимодействие человека с машиной или машины с машиной. Координация и сотрудничество между людьми и машинами требуют четко определенных интерфейсов; помимо всего прочего, необходимо уяснить, кто к кому должен адаптироваться – машина к человеку или, наоборот, человек к машине. Безопасность данных и права собственности также чрезвычайно важны и должны быть защищены законом. Требуется и четкая правовая экспликация проблематики ответственности в таком человеко-машинном взаимодействии.

Результатом этих изменений должна стать всеобъемлющая трансформация универсума трудовых отношений [Bömer, Kehl, Nierling, 2017], требующая не только новых инициатив по организации обучения на ранней стадии навыков работы в цифровой среде и последующей переподготовке, но и оказывающая все более сильное воздействие на безопасность внеофисного труда с гибким графиком (краудворкинг), будущую роль профсоюзов или новых организаций, представляющих интересы работников цифровой эпохи, международное разделение труда на глобальном рынке занятости, развитие систем социального обеспечения. В свете всех этих стратегически значимых воздействий на экономику и социальную сферу насущной необходимостью становятся инновационная активность и проектирование новых технологий, включающие учет аспектов социальной ответственности, а также раннее предупреждение и – насколько возможно – предотвращение негативных

социальных и экологических эффектов. Таким образом, социальная оценка техники оказывается востребованной и в условиях Четвертой промышленной революции. Дилемма Коллингриджа, будучи явной идеализацией, указывает на существенные трудности, но не на невозможность или бесполезность ТА в условиях цифровизации.

### **Беспилотный автомобиль**

Будущее автовождения тесно связано с продолжающейся цифровой революцией. В последние годы человек за рулем пользуется постоянно возрастающим объемом помощи со стороны передовых сенсорных устройств, осуществляющих оценку собранных данных в режиме реального времени, и исполнительных механизмов, выполняющих действия на основе заданных алгоритмов. Процессоры и датчики все в большей степени контролируют ситуацию на дороге и определяют последующие действия, позволяющие эффективно управлять автомобилем с учетом конкретных условий трафика. Эта тенденция уже привела к частичной автоматизации автовождения. Высокоавтоматизированные системы могут автономно менять полосу движения и выполнять другие функции без вмешательства человека за рулем. В некоторых странах созданы испытательные полигоны, на которых могут эксплуатироваться высоко- или полностью автоматизированные транспортные средства.

Необходимо признать, что беспилотное автовождение открывает широкий спектр новых возможностей мобильности, выходящих за рамки традиционной индивидуальной мобильности, связанной с использованием частного автотранспорта: меняются концептуальные подходы к передвижению на автомашинах, формируются новые бизнес-модели для поставщиков беспилотных авто и сопутствующих услуг, возникают неожиданные комбинации частного и общественного транспорта вплоть до возможного стирания традиционных границ между ними. Одно это обстоятельство делает беспилотные автомобили, возможно, прорывной инновацией. Но, несмотря на многочисленные и далеко идущие положительные эффекты внедрения беспилотного автомобиля в плане безопасности и комфорта, не следует забывать, что оно влечет за собой и определенные риски. Некоторые из них хорошо известны по практике традиционного автовождения, тогда как другие обнаруживают себя лишь сейчас, прежде всего, в связи с нарастающими масштабами цифровизации автовождения [Maurer et al., 2016].

В частности, продолжаются исследования коэволюционной динамики цифровизации и доминирующих паттернов мобильности. Распространение различных инкарнаций автоматизированных транспортных средств (АТС) будет способствовать увеличению сложности дорожного движения в зависимости от того, как АТС будут взаимодействовать друг с другом, с транспортными средствами, управляемыми людьми, и с пешеходами. Поскольку дорожное движение детерминируется сетью постоянно согласуемых и переопределяемых социальных отношений, которые выходят далеко за рамки простого подчинения правилам поведения, беспилотные автомобили на определенном уровне автоматизации становятся социальными акторами в этой сети. Это

ставит многочисленные исследовательские проблемы, например, будет ли беспилотникам атрибутироваться агентность в дорожном трафике, должно ли совместимое с автоматизацией поведение людей на дорогах выступать в качестве регулирующего принципа, могут ли ценности быть реализованы в дизайне беспилотных автомобилей, и если да, то какие и как [Grunwald, 2018].

Изменение отношений ответственности потребует особого рассмотрения как с этической, так и с юридической точек зрения. Принципиальный подход был сформулирован в докладе Комиссии по этике Федерального министерства транспорта и цифровой инфраструктуры Германии: «В случае беспилотных систем человеко-машинный интерфейс должен быть спроектирован таким образом, чтобы в любой момент времени можно было со всей очевидностью определить, на ком лежит индивидуальная ответственность, особенно ответственность за управление. Распределение ответственности (и, следовательно, подотчетности), например, в отношении времени и условий доступа, должно быть задокументировано и сохранено. В особенности это относится к процедурам перехода ответственности от человека к технологии» [Ethics Commission, 2017]. Решение задач, связанных с определением и вменением ответственности, относится ко всем формам вождения с высоко- или частично автоматизированной системой управления, тогда как полностью автоматизированное или беспилотное вождение здесь не является проблемой, поскольку управление полностью обеспечивается технологиями.

Еще одна дискуссионная тема связана с ситуациями безнадёжного выбора, когда, например, системе управления беспилотным автомобилем предстоит решать, совершить ли наезд на перебегающих дорогу в неположенном месте детей или, уклоняясь от столкновения на дороге, сбить идущую по соседнему тротуару пожилую пару. До сих пор такие ситуации рассматривались как трагические события, когда кто-то оказался не в том месте и не в то время. Однако теперь, когда речь идет о широком распространении беспилотного автотранспорта, различные вариации такого рода дилеммы приобретают совершенно новое и в первую очередь практическое значение. В конечном счете тот или иной вариант действий в подобной ситуации должен быть заложен в алгоритм управляющего устройства, и, таким образом, невозможный с этической точки зрения выбор в конечном счете должны будут сделать люди, отвечающие за программные решения [Wallach, Allen, 2009]. Ответственность, таким образом, снова возвращается в сферу человеческой агентности, но уже в этой сфере сохраняется множество вопросов о конкретных субъектах ответственности, способах ее определения и вменения.

### **Роботизированные системы для ухода за тяжелобольными и престарелыми**

Глобальный стресс 2020–2021 гг. – пандемия коронавируса COVID 19 – выявил насущную потребность в осознании непосредственной и, очевидно, неустранимой связи между процессами цифровизации и проблемами здравоохранения, а в конечном счете – безопасности и устойчивого развития социума. Еще на пике первой волны пандемии во всем мире начался полномасштабный

разворот к изучению комплексных эффектов использования цифровых технологий для реализации антиэпидемических мер, обеспечения занятости, трансформации форм деловой коммуникации и т.д. Можно сказать, что рабочая повестка социальной оценки техники в отношении цифровизации в силу чрезвычайных обстоятельств оказалась в центре внимания всего научного сообщества, политических и деловых кругов, гражданского общества. Однако и вне контекста пандемии цифровая трансформация в сфере здравоохранения является темой большой важности. Остановимся на одном из направлений – использовании роботизированных систем для ухода за тяжелобольными и престарелыми.

В случае разработки технологий ухода на основе искусственного интеллекта подход конструктивной оценки техники оказывается весьма востребованным. С учетом тенденции старения общества и увеличения доли людей с ограниченными физическими и когнитивными способностями в общей численности населения уход за больными и престарелыми становится все более серьезной социальной проблемой. По большинству оценок, автономно функционирующие роботы для обслуживания и ухода за больными и престарелыми, а также вспомогательные технологии в сочетании с нейротехнологиями (например, экзоскелетами) имеют большой потенциал для решения данной проблемы. ИИ играет при этом важную роль, позволяя технически автономным системам адекватно действовать в сложных условиях, выявляя в режиме реального времени чувствительные изменения состояния пациентов, например, в случае вариабельной деменции [Decker et al., 2017]. Своевременное распознавание таких изменений чрезвычайно важно, поскольку интенсивность контакта между пациентами и соответствующими техническими устройствами очень высока, и любой технический сбой может иметь серьезные или даже фатальные последствия. Использование роботов и других технологий ухода за тяжелобольными и престарелыми сопряжено с широким спектром вопросов человеческого достоинства, этики и права, которые необходимо в полной мере учитывать в процессе проектирования. На стадии проектирования должны быть приняты во внимание и практические потребности пациентов, их родственников и медицинского персонала, чтобы в дальнейшем предотвратить возникновение нежелательных эффектов и избежать переналадки или досрочного вывода из эксплуатации роботизированных систем. Здесь, таким образом, решающее значение имеет не только подход к разработке технологий, ориентированный на проблему и потребителя, но и оптимальный выбор момента времени. Это означает, что проектирование роботизированных систем для медицинского ухода нельзя доверить только программистам, медикам, юристам и специалистам по этике – в данном случае требуется создание и устойчивое функционирование трансдисциплинарной сети, включающей помимо профильных специалистов также и представителей тех, для кого наличие подобных систем имеет жизненно важное значение.

Аналогичный подход требуется и для разработки многих других систем, обеспечивающих цифровую трансформацию. Помимо технических знаний и мастерства в процессе проектирования цифровых технологий необходимы информированность о вероятных воздействиях, этическая чувствительность



и юридическая компетентность, требующие осведомленности в соответствующих дисциплинарных областях. Учет в процессе разработки цифровых технологий социальных ценностей предполагает включение в этот процесс нескольких итераций взаимодействия с основными группами пользователей, а также тех, на кого проектируемые технологии могут оказать существенное влияние. В этом взаимодействии моральные и ценностные аспекты применения новой технологии должны быть эксплицированы таким образом, чтобы определить дополнительные требования к проектированию, внести коррективы, позволяющие предотвратить конфликты ценностей и интересов.

Но даже если процессы цифровой трансформации разворачиваются благодаря серии сугубо технократических решений, принимаемых в кабинетах чиновников на основе недостаточно репрезентативной экспертизы и без участия заинтересованных сторон, то и в этом случае конструктивистская парадигма социальной оценки техники позволяет рассматривать такие решения, например, многие модели электронного правительства или программы перехода к алгоритмическому управлению городским хозяйством в мегаполисах, как примеры «социотехнического воображаемого» [Jasanoff, Kim, 2015]. Даже самые «забюрокраченные» программы цифровизации основываются на определенных нормативных образах экономического, социального и технологического порядка. Подобные программы обычно исходят из упрощенных детерминистских установок, порождающих неадекватные ожидания того, что внедрение цифровых технологий само по себе резко повысит качество управления и обеспечит эффективную обратную связь с гражданами. Можно говорить даже о наличии специфической мифологии [Curran, Fenton, Freeman, 2016], убеждающей аудиторию в том, что один лишь переход к полномасштабному использованию Big Data, Интернета вещей, геоинформационных систем и т.п. приведет к кардинальной трансформации классической «веберовской» системы государственного управления [Трахтенберг, 2017] либо – тем более – управленческой традиции, остающейся весьма далекой от идеала рациональной бюрократии. Неудивительно, что немалая часть бюрократических проектов цифровизации терпит полный или частичный провал.

### **Институциональные эффекты цифровизации**

Отдельная проблема – это институциональные эффекты цифровизации, в частности то, что программное обеспечение, алгоритмы, сети и услуги на основе ИИ начинают выступать в качестве квазиинститутов, оказывая организующее, стимулирующее либо ограничивающее воздействие на социальное поведение. Китайская система социального кредита, предлагающая широкий набор стимулов «хорошего» социального поведения и санкций в отношении «плохого» – наиболее яркий пример такого рода воздействия. Другими примерами институциональных эффектов сложных алгоритмов являются правила интернет-фильтров, системы управления цифровыми правами, интернет-архитектуры и протоколы, поисковые системы, системы электронной торговли, социальные сети и онлайн-сообщества, системы ИИ, правила принятия решений в аналитике больших данных и т.д. [Orwat et al., 2010]. Постепенный переход

межличностных взаимодействий в цифровые системы означает, что некий набор социальных правил и базовых ценностей в эти системы «запрограммирован». Тем самым применение правил автоматизируется, причем этот процесс идет в направлении усиливающейся детализации.

В исследованиях, посвященных институциональным эффектам программного обеспечения [van den Hoven, 2007; Brey, 2009], обращается внимание на то, что регулирование с использованием программного обеспечения имеет тенденцию «перезаписывать» или ставить под угрозу традиционные правила, ценности и общественные ожидания. Со временем происходит инфильтрация подобного регулирования в системы публичного права, хотя эти новации чаще всего не имеют демократической легитимности, но инспирируются транснациональными корпорациями.

Институциональные эффекты программного обеспечения служат еще одним аргументом в пользу того, чтобы проблемы морали и ответственности рассматривались на максимально ранней стадии деятельности, результатом которой становится появление и внедрение новой технологии. Дилемма Коллингриджа указывает на то, что предпочтительно начинать действовать в этом направлении еще на стадии исследований и инновационных разработок. Свое концептуальное оформление эта установка получила в рамках дискуссии об «ответственных исследованиях и инновациях» (Responsible Research and Innovations / RRI) [Owen, Bessant, Heintz, 2013], которая привела к своеобразному синтезу идей социальной оценки техники, инженерной этики и наработок междисциплинарной исследовательской программы «Наука, техника и общество» (Science, Technology and Society / STS).

### **Ответственность в исследованиях и инновациях (RRI)**

Появление концепции RRI отражает понимание того, что прежние подходы, несмотря на несомненное повышение рефлексивности научно-технической деятельности, далеко не оправдывают всех надежд на гармоничный прогресс технологий и общества. Отметим основные особенности RRI:

– Идея «формирования инноваций» дополняет или даже заменяет прежнюю конструктивистскую установку на «формирование технологий». Этот сдвиг отражает понимание того, что не технологии как таковые влияют на общество, а именно инновации, посредством которых обеспечивается взаимодействие техники и общества. Следовательно, целенаправленное воздействие, учитывающее потребности, ожидания и ценности общества, должно оказываться на стадии инновационной активности.

– В связи с этим необходимо уделять пристальное внимание социальным контекстам науки и новых технологий. «Ответственные инновации» имеет смысл рассматривать как еще один шаг на пути в направлении социально формируемого спроса на результаты инновационной деятельности, обеспечивающего совместимость этих результатов с доминирующими в обществе ценностями.

– Вместо установок классической научной парадигмы на дистанционное наблюдение концепция RRI предлагает активное вмешательство в процессы исследований, разработок и инноваций.

– Новые формы управления научно-технической деятельностью должны привести к еще большему расширению диапазона заинтересованных акторов, участвующих в диалоге и принятии решений относительно конкретных задач и направлений этой деятельности [Grunwald, 2011].

Все отмеченные особенности RRI показывают, что это движение представляет собой радикальную версию постнормальной науки [Funtowicz, Ravetz, 1993], феномен которой означает не просто сближение научного познания и социальной практики, но понимание научно-технической деятельности как процесса, обусловленного контекстом применения знания, представлениями о социальных потребностях и потенциальных потребителях. RRI требует более интенсивного меж- и трансдисциплинарного сотрудничества между инженерными, информационными, социальными науками, исследованиями управления и прикладной этикой.

То обстоятельство, что цифровые технологии открывают множество новых возможностей практически во всех сферах жизни современного общества и одновременно создают новые проблемы, вплоть до экзистенциальных, делает их приоритетной сферой для RRI и социальной оценки техники. Практический эффект социальной оценки цифровых технологий и применения к ним подхода RRI во многом связан с институциональными особенностями управления научно-техническим развитием в той или иной стране (наднациональном интеграционном объединении). Свою роль при этом играют и особенности политической культуры, моделей участия представителей гражданского общества в решении проблем, связанных с научно-технической и инновационной активностью, различия систем ценностей тех или иных обществ. Можно предположить, что вызовы цифровой экономики и цифрового общества будут способствовать трансформации устоявшихся подходов в сферах научно-технической политики и ее экспертного обеспечения. Весьма вероятно, что ключевой темой социальной оценки техники в контексте цифровизации станут экзистенциальные и этические проблемы человеко-машинного взаимодействия, связанные с тем, что искусственный интеллект во все возрастающей степени будет «осваивать» те сферы когнитивной деятельности, которые наиболее значимы для человеческой идентичности, включая и эмоциональные переживания. В то же время в центре внимания будет оставаться гигантское преобразующее воздействие цифровизации на критические инфраструктуры, такие как энергосистемы, беспилотный транспорт, автоматизированные производства и т.д. Здесь должен рассматриваться весь комплекс воздействий, включая социальную устойчивость и уязвимость, стратегическое управление и контроль технологий, в конечном счете – фундаментальные основы существования индивидов и групп в условиях цифрового общества.

Наконец, следует исходить из того, что основные эффекты цифровизации относятся к числу глобальных проблем, сопоставимых в среднесрочной перспективе с изменениями климата или сокращением биологического разнообразия. Тот факт, что цифровая среда становится сферой острой геополитической конкуренции, лишь подчеркивает настоятельную потребность в достижении хотя бы минимально необходимого уровня международного сотрудничества,

позволяющего выработать базовые правила и ограничения в глобальных процессах цифрового перехода. Есть основания опасаться, что различия национальных интересов и стремление государственных акторов к сохранению суверенитета и в цифровой реальности приведут к тому, что время для необходимых скоординированных усилий международного сообщества, направленных на регулирование цифровых технологий, окажется упущенным. Актуальная дискуссия об антропоцене [Crutzen, Stoermer, 2000; Trischler, 2016] свое дальнейшее развитие, очевидно, получит в контексте цифровизации и ее комплексных эффектов. Цифровизация становится триггером и индикатором быстрых изменений в мировой экономике, материальных и энергетических потоках, преобразовании социальных потребностей, поведения отдельных индивидов и крупных сообществ. В плане воздействия на глобальное экологическое равновесие цифровизация оказывается амбивалентной: цифровые технологии вносят все возрастающий вклад в противоречащую целям устойчивого развития динамику эксплуатации ресурсов и производства отходов, но одновременно именно с цифровыми технологиями связаны возможности решения этих проблем. Иначе говоря, цифровизация – это процесс с открытым финалом. Добиться его переориентации на обеспечение устойчивого развития человеческой цивилизации и возвращение к равновесному состоянию важнейших экосистем планеты – ключевая проблема глобальной ответственности политических лидеров, бизнеса, ученых и гражданского общества.

### Список литературы

- Глазьев, Харитонов, 2009 – Глазьев С.Ю., Харитонов В.В. (ред.). Нанотехнологии как ключевой фактор нового технологического уклада в экономике. М.: Травант, 2009. 304 с.
- Горохов, 2015 – Горохов В.Г. Оценка техники как прикладная философия техники и новая научно-техническая дисциплина // Гений Шухова и современная эпоха. Материалы Международного конгресса. М.: Изд-во МБГТУ им. Н.Э. Баумана, 2015. С. 241–249.
- Ефременко, 2002 – Ефременко Д.В. Введение в оценку техники. М.: Изд-во Международного независимого эколого-политологического ун-та, 2002. 186 с.
- Трахтенберг, 2017 – Трахтенберг А.Д. Идеологический концепт электронного правительства: как работает риторика разрыва? // Научный ежегодник Института философии и права Уральского отделения Российской академии наук. 2017. Т. 17. Вып. 2. С. 41–58.
- Шваб, 2017 – Шваб К. Четвертая промышленная революция. М.: Эксмо, 2017. 288 с.
- Bijker, Hughes, Pinch, 1987 – Bijker W., Hughes T., Pinch T. (eds). The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology. Cambridge – London: MIT Press, 1987. 482 p.
- Börner, Kehl, Nierling, 2017 – Börner F., Kehl C., Nierling L. Chancen und Risiken mobiler und digitaler Kommunikation in der Arbeitswelt // TAB-Arbeitsbericht No. 174. Berlin: Büro für Technikfolgenabschätzung beim Deutschen Bundestag, 2017. 243 S.
- Brey, 2009 – Brey P. Values in Technology and Disclosive Computer Ethics / Floridi L. (ed.). The Cambridge Handbook of Information and Computer Ethics. Cambridge: Cambridge University Press, 2009. P. 41–58.
- Brown, Marsden, 2013 – Brown I., Marsden C.T. Regulating Code: Good Governance and Better Regulation in the Information Age (Information Revolution and Global Politics). Cambridge, MA: MIT Press, 2013. 288 p.

Collingridge, 1980 – *Collingridge D.* The Social Control of Technology. New York: St. Martin's Press, 1980. 200 p.

Collingridge, 1990 – *Collingridge D.* Technology Organizations and Incrementalism: The Space Shuttle // *Technology Analysis & Strategic Management*. 1990. Vol. 2. No. 2. P. 181–200.

Crutzen, Stoermer, 2000 – *Crutzen P.J., Stoermer E.F.* The “Anthropocene” // *Global Change Newsletter*. 2000, May. Vol. 41. P. 17–18.

Curran, Fenton, Freeman, 2016 – *Curran J., Fenton N., Freeman D.* Misunderstanding the Internet. London: Routledge, 2016. 234 p.

Decker et al., 2017 – *Decker M., Weinberger N., Krings B., Hirsch J.* Imagined Technology Futures in Demand-oriented Technology Assessment // *Journal of Responsible Innovation*. 2017. Vol. 4. No. 2. P. 177–196.

Ethics Commission, 2017 – *Ethics Commission*. Automated and Connected Driving. Final Report. Berlin: Federal Ministry of Transportation and Digital Infrastructures, 2017. URL: <https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.html> (дата обращения: 19.05.2021).

Funtowicz, Ravetz, 1993 – *Funtowicz S.O., Ravetz J.* Science in the Post-Normal Age // *Futures*. 1993. Vol. 25. No. 7. P. 739–756.

Grunwald, 2011 – *Grunwald A.* Responsible Innovation: Bringing Together Technology Assessment, Applied Ethics, and STS Research // *Enterprise and Work Innovation Studies*. 2011. No. 7. P. 9–31.

Grunwald, 2018 – *Grunwald A.* Self-Driving Cars: Risk Constellation and Acceptance Issues // *DELPHI-Interdisciplinary Review of Emerging Technologies*. 2018. No. 1. P. 8–13.

Grunwald, 2019 – *Grunwald A.* Technology Assessment in Practice and Theory. London: Routledge, 2019. 274 p.

Hermann, Pentek, Otto, 2016 – *Hermann M., Pentek T., Otto B.* Design Principles for Industry 4.0 Scenarios // *Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS)*. Washington, DC: IEEE Computer Society: 2016. P. 3928–3937.

van den Hoven, 2007 – *van den Hoven J.* ICT and Value Sensitive Design // *Goujon P., Lavelle S., Duquenoy P., Kimppa K., Laurent V.* (eds.) *The Information Society: Innovations, Legitimacy, Ethics and Democracy*. Boston: Springer, 2007. P. 67–72.

Jasanoff, Kim, 2015 – *Jasanoff S., Kim S.-H.* (eds.) *Dreamscapes of Modernity. Socio-technical Imaginaries and the Fabrication of Power*. Chicago: University of Chicago Press, 2015. 360 p.

Kagermann, Lukas, Wahlster, 2011 – *Kagermann H., Lukas W.-D., Wahlster W.* Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution // *VDI Nachrichten*. Berlin, 01.04.2011. URL: [http://www.wolfgang-wahlster.de/wordpress/wp-content/uploads/Industrie\\_4\\_0\\_Mit\\_dem\\_Internet\\_der\\_Dinge\\_auf\\_dem\\_Weg\\_zur\\_vierten\\_industriellen\\_Revolution\\_2.pdf](http://www.wolfgang-wahlster.de/wordpress/wp-content/uploads/Industrie_4_0_Mit_dem_Internet_der_Dinge_auf_dem_Weg_zur_vierten_industriellen_Revolution_2.pdf) (дата обращения: 08.04.2021).

Maurer et al., 2016 – *Maurer M., Gerdes J., Lenz B., Winner H.* (eds.) *Autonomous driving. Technical, Legal and Social aspects*. Heidelberg: Springer Open, 2016. 698 p.

Orwat et al., 2010 – *Orwat C., Raabe O., Buchmann E., Anandasivam A., Freytag J.-C., Helberger N., Ishii K., Lutterbeck B., Neumann D., Otter T., Pallas F., Reussner R., Sester P., Weber K., Werle R.* Software als Institution und ihre Gestaltbarkeit // *Informatik-Spektrum*. Bd. 33. 2010. No. 6. S. 626–633.

Owen, Bessant, Heintz, 2013 – *Owen R., Bessant J., Heintz M.* (eds.) *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*. London: Wiley, 2013. 306 p.

Rip, Misa, Schot, 1995 – *Rip A., Misa T., Schot J.* (eds.) *Managing Technology in Society*. London: Pinter, 1995. 361 p.

Trischler, 2016 – *Trischler H.* The Anthropocene. A Challenge for the History of Science, Technology, and the Environment // *NTM Zeitschrift für Geschichte der Wissenschaften, Technik und Medizin.* 2016. Bd. 24. S. 309–335.

Wallach, Allen, 2009 – *Wallach W., Allen C.* Moral Machines: Teaching Robots Right from Wrong. New York: Oxford University Press, 2009. 273 p.

## Digital transformation and technology assessment

*Armin Grunwald*

Institute of Technology Assessment and Systems Analysis, Karlsruhe Institute of Technology, P.O. Box 3640, Karlsruhe, 76021, Germany, e-mail: armin.grunwald@kit.edu

*Dmitry V. Efremenko*

Institute of Scientific Information for Social Sciences, Russian Academy of Sciences. 51/21 Nakhimovskiy av., Moscow, 117997, Russian Federation; e-mail: efdv2015@mail.ru

The article discusses the possibilities of using the technology assessment as a special area of complex interdisciplinary research of the long-term consequences of the development of new technologies in the context of the digital transformation of modern society. Assessment of digital technologies allows supporting the political decision-making process, as well as finding options for resolving conflicts of social actors caused by digitalization processes. At the same time, it is necessary to take into account the institutional effects of the development of digital technologies. Integration of the principles of ethics and social responsibility into the process of social control of digital transformation is of fundamental importance. Particular attention is paid to such areas of digital transformation as the development of Industry 4.0, the widespread introduction of autonomous driving and care robots.

**Keywords:** digital transformation, technology assessment, industry 4.0, autonomous driving, care robots, responsible research and innovations

## References

Bijker, W., Hughes, T., Pinch, T. (eds). *The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology.* Cambridge; London: MIT Press, 1987. 482 pp.

Börner, F., Kehl, C., Nierling, L. “Chancen und Risiken mobiler und digitaler Kommunikation in der Arbeitswelt”, *TAB-Arbeitsbericht Nr. 174.* Berlin: Büro für Technikfolgenabschätzung beim Deutschen Bundestag, 2017. 243 S.

Brey, P. “Values in Technology and Disclosive Computer Ethics”, *The Cambridge Handbook of Information and Computer Ethics*, Floridi L. (ed.). Cambridge: Cambridge University Press, 2009, pp. 41–58.

Brown, I. & Marsden, C.T. *Regulating Code: Good Governance and Better Regulation in the Information Age (Information Revolution and Global Politics).* Cambridge, MA: MIT Press, 2013. 288 pp.

Collingridge, D. *The Social Control of Technology.* New York: St. Martin’s Press, 1980. 200 pp.

Collingridge D. “Technology Organizations and Incrementalism: The Space Shuttle”, *Technology Analysis & Strategic Management*, 1990, vol. 2, no. 2, pp. 181–200.

Crutzen, P.J. & Stoermer, E.F. The “Anthropocene”, *Global Change Newsletter*, 2000, May, vol. 41, pp. 17–18.

Curran, J., Fenton, N., Freeman, D. *Misunderstanding the Internet*. London: Routledge, 2016. 234 pp.

Decker, M., Weinberger, N., Krings, B., Hirsch, J. “Imagined Technology Futures in Demand-oriented Technology Assessment”, *Journal of Responsible Innovation*, 2017, vol. 4, no. 2, pp. 177–196.

Efremenko, D.V. *Vvedenie v otsenku tehniki* [Introduction to Technology Assessment]. Moscow: Izdatel'stvo Mezhdunarodnogo nezavisimogo ekologo-politologicheskogo universiteta, 2002. 186 pp. (In Russian)

Ethics Commission. *Automated and Connected Driving*. Final Report. Berlin: Federal Ministry of Transportation and Digital Infrastructures, 2017. [<https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.html>, accessed on 19.05.2021].

Funtowicz, S.O. & Ravetz, J. “Science in the Post-Normal Age”, *Futures*, 1993, vol. 25, no. 7, pp. 739–756.

Glazyev, S.Yu., Kharitonov, V.V. (eds.) *Nanotechnologii kak klyuchevoy factor novogo technologicheskogo uklada v ekonomike* [Nanotechnology as a key factor in the new technological order in the economy]. Moscow: Trovant Publ., 2009. 304 pp. (In Russian).

Gorokhov, V.G. “Otsenka tehniki kak prikladnaya filosofia tehniki i novaya nauchno-technicheskaya disciplina” [Technology assessment as an applied philosophy of technology and a new scientific and technical discipline], *Genij Schukhova i sovremennaya epokha. Materialy mezhdunarodnogo kongressa*. Moscow: Izdatel'stvo MBGTU im. N.E. Baumana, 2015, pp. 241–249. (In Russian)

Grunwald, A. “Responsible Innovation: Bringing Together Technology Assessment, Applied Ethics, and STS Research”, *Enterprise and Work Innovation Studies*, 2011, no. 7, pp. 9–31.

Grunwald, A. “Self-Driving Cars: Risk Constellation and Acceptance Issues”, *DELPHI-Interdisciplinary Review of Emerging Technologies*, 2018, no. 1, pp. 8–13.

Grunwald, A. *Technology Assessment in Practice and Theory*. London: Routledge, 2019. 274 pp.

Hermann, M., Pentek, T., Otto, B. “Design Principles for Industry 4.0 Scenarios”, *Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS)*. Washington, DC: IEEE Computer Society, 2016, pp. 3928–3937.

van den Hoven, J. “ICT and Value Sensitive Design”, *The Information Society: Innovations, Legitimacy, Ethics and Democracy*, Goujon, P., Lavelle, S., Duquenoy, P., Kimppa, K., Laurent, V. (eds.). Boston: Springer, 2007, pp. 67–72.

Jasanoff, S., Kim, S.-H. (eds.) *Dreamscapes of Modernity. Sociotechnical Imaginaries and the Fabrication of Power*. Chicago: University of Chicago Press, 2015. 360 pp.

Kagermann, H., Lukas, W.-D., Wahlster, W. “Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution”, *VDI Nachrichten*. Berlin, 01.04.2011. [[http://www.wolfgang-wahlster.de/wordpress/wp-content/uploads/Industrie\\_4\\_0\\_Mit\\_dem\\_Internet\\_der\\_Dinge\\_auf\\_dem\\_Weg\\_zur\\_vierten\\_industriellen\\_Revolution\\_2.pdf](http://www.wolfgang-wahlster.de/wordpress/wp-content/uploads/Industrie_4_0_Mit_dem_Internet_der_Dinge_auf_dem_Weg_zur_vierten_industriellen_Revolution_2.pdf), accessed on 08.04.2021].

Maurer, M., Gerdes, J., Lenz, B., Winner, H. (eds.) *Autonomous driving. Technical, Legal and Social aspects*. Heidelberg: Springer Open, 2016. 698 pp.

Orwat, C., Raabe, O., Buchmann, E., Anandasivam, A., Freytag, J.-C., Helberger, N., Ishii, K., Lutterbeck, B., Neumann, D., Otter, T., Pallas, F., Reussner, R., Sester, P., Weber, K., Werle, R. “Software als Institution und ihre Gestaltbarkeit”, *Informatik-Spektrum*, 2010, bd. 33, no. 6, S. 626–633.

Owen, R., Bessant, J., Heintz, M. (eds.) *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*. London: Wiley, 2013. 306 pp.

Rip, A., Misa, T., Schot, J. (eds). *Managing Technology in Society*. London: Pinter, 1995. 361 pp.

Schwab, K. *Chetvertaya promyshlennaya revolyutsia* [Fourth industrial revolution]. Moscow: Eksmo Publ., 2017. 288 pp. (In Russian)

Trachtenberg, A.D. "Ideologicheskij koncept elektronogo pravitelstva: kak rabotaet ritorika razryva?" [Ideological concept of e-government: how does the rhetoric of the breaking work?], *Nauchnyj ezhegodnik Instituta filosofii i prava Uralskogo otdelenia Rossijskoj Akademii nauk*, 2017, vol. 17, iss. 2, pp. 41–58. (In Russian)

Trischler, H. "The Anthropocene. A Challenge for the History of Science, Technology, and the Environment", *NTM Zeitschrift für Geschichte der Wissenschaften, Technik und Medizin*, 2016, bd. 24, S. 309–335.

Wallach, W., Allen, C. *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press, 2009. 273 pp.