



Три ответа на антропоморфизм в социальной робототехнике: на пути к критическому, реляционному и герменевтическому подходу

Марк Кокельберг¹Принято: 23 февраля 2021 ©
Автор(ы) 2021

Абстрактный

И дизайнеры, и пользователи социальных роботов склонны антропоморфизировать роботов. Сосредоточив внимание на вопросе, как концептуализировать отношения между роботами и людьми, в этой первой статье излагаются два противоположных философских взгляда на эти отношения, которые связаны с различными нормативными реакциями на антропоморфизм и антропоморфизацию. Затем он приводит доводы в пользу третьей точки зрения: переходя между тем, что он называет «наивным инструментализмом» и «некритическим постгуманизмом», он развивает герменевтический, реляционный и критический подход. Парадоксально, но, раскрывая человеческое измерение робототехники в ее использовании и развитии, эта точка зрения позволяет критически обсудить антропоморфизацию роботов. В то же время, и это опять же несколько парадоксально, оно избегает позиции наивного инструменталиста, серьезно воспринимая роль роботов как инструмента в более широкой мошеннической технологии. Таким образом, третья точка зрения ставит под сомнение дуализм, принятый в ходе дебатов. Затем в статье исследуется, что это означает для области социальной робототехники и образования ученых-компьютерщиков и инженеров. Он предлагает реформу, основанную на реляционном понимании самой области, и предлагает предложения по роли граждан-пользователей.

Ключевые слова Социальная робототехника Отношения человека и робота Антропоморфизм Антропоморфизация Инструментализм Постгуманизм · Герменевтика · Реляционный подход · Власть

1 Введение: антропоморфизм и антропоморфизация в социальных сетях Робототехника

Социальные робототехники разрабатывают роботов, которые должны функционировать более естественно в социальных ситуациях, чем машины из прошлого. Для этой цели их моделируют по образцу людей и других высших приматов, которых, в свою очередь, склонны понимать посредством метафоры своего рода социальной машины, хотя и развивающейся и эволюционировавшей. С помощью литературы по психологии, когнитивной науке, нейробиологии и другим областям исследователи в области социальной робототехники и взаимодействия человека и робота (HRI) определяют функции, которые считаются необходимыми для социального взаимодействия, и затем реализуют их в работе. Например, Даутенхан ссылается на работу в области HRI (в частности [1]) и на гипотезу социального мозга, которая утверждает, что интеллект приматов развился из-за необходимости иметь дело со все более сложными социальными задачами.

динамики, чтобы определить ряд характеристик социального взаимодействия, реализуемых в социальных роботах. Она утверждает, что «социально интерактивные роботы обладают следующими характеристиками: выражать и/или воспринимать эмоции; общаться с помощью диалога на высоком уровне; изучать модели или распознавать других агентов; устанавливать и/или поддерживать социальные отношения; используйте естественные сигналы (взгляд, жесты и т. д.); проявлять отличительную индивидуальность и характер; и может учиться и/или развивать социальные компетенции». [2:684; 1:145] Реализация таких социальных функций часто приводит к тому, что робот демонстрирует антропоморфный внешний вид и поведение. Однако цель не в том, чтобы создать человекоподобных роботов как таковых (или роботов, похожих на приматов, если уж на то пошло). Антропоморфизация в этом контексте преследует более прагматичную цель. В социальной робототехнике это стратегия интеграции роботов в социальную среду человека. Это средство дать роботам «способность участвовать в значимом социальном взаимодействии с людьми» [3:178]. И дизайнеры, и пользователи затем склонны антропоморфизировать таких роботов при взаимодействии с ними, приписывая им антропоморфные характеристики, такие как индивидуальность, живость и так далее. На следующих страницах я буду называть «антропоморфизм», «антропоморфизацию» и «антропоморфизацию» соответственно явлением и деятельностью по приписыванию роботам человеческих х

* Марк Кокельберг
mark.coeckelbergh@univie.ac.at

¹ Венский университет, Вена, Австрия

(пользователями при взаимодействии с роботами) и созданием таких характеристик дизайнерами и разработчиками роботов.

Иногда антропоморфизм не подразумевается. Социальные робототехники знают об этом и поэтому пытаются контролировать это явление. Своим дизайном они могут пытаться поощрять или препятствовать антропоморфизации со стороны пользователя. Например, некоторые робототехники намеренно пытаются создать роботов, максимально похожих на человека: Исигуро [4] утверждал, что гуманоиды являются идеальными интерфейсами. Другие строят менее человекоподобно роботы, например роботы, похожие на животных, или вообще стараются избегать антропоморфизма. Теоретически антропоморфизм не кажется необходимым для социальной робототехники; Вполне возможно, что робот будет обладать упомянутыми выше характеристиками, но не будет похож на человека. Но на практике многие специалисты по социальной робототехнике приветствуют некоторую степень или форму антропоморфизма, поскольку, как утверждается, он облегчает взаимодействие между людьми и роботами. Сегодня антропоморфизм является одним из факторов, которые измеряются при взаимодействии молодого человека с роботом (HRI) [5]. Признано, что антропоморфизм представляет собой проблему, но в конечном итоге цель состоит в том, чтобы « содействовать интеграции человекоподобных машин в реальный мир » [6]. антропоморфизм помогает в достижении этой цели.

Роботы очень хорошо подходят для этой цели [7]: из-за своего внешнего вида и присутствия они обладают высокой « антропоморфизируемостью » [6]. Если уже в 1990-е годы Ривз и Насс обнаружили, что люди относятся к компьютерам как к реальным людям [8], с роботами это работает еще лучше, особенно с роботами-гуманоидами [9]. Но на антропоморфизм влияет не только внешний вид: роботы могут предлагать всевозможные сигналы во время взаимодействия [10], существуют и другие факторы, такие как автономность робота, предсказуемость и т. д. Однако для целей данной статьи интересно, что антропоморфизм является не только результатом того, как робот выглядит или что он делает, но также зависит от таких характеристик наблюдателя, как социальное происхождение и пол [6]. Например, люди применяют к роботам социальные категории, такие как членство в группах.

[11] Это означает, что социальный аспект может помочь лучше понять и, возможно, также лучше оценить антропоморфизм во взаимодействии человека и робота. Это также предполагает, что наряду с объективными характеристиками робота субъективность

Пользователь имеет значение.

Психологи помогают робототехникам в этом проекте использования антропоморфизации с целью разработки социальных роботов, проводя эмпирические исследования восприятия роботов (в зависимости от их особенностей) — например, разрабатывая вопросы и анализируя результаты с целью изучения приемлемости [12, 13]. . Робототехники также импортируют психологическую теорию человеческой социальности и проводят исследования в междисциплинарном контексте: они используют ключевые парадигмы психологии, а также используют и воспроизводят результаты взаимодействия человека с человеком; часто встречается экспериментальные подходы [14]. Помимо формирования конструктивных особенностей робота, каркас также является методом

Предложите людям антропоморфизировать робототехнические технологии: роботам можно дать имя или предысторию, что повлияет на реакцию людей на роботов [15]. Опять же, антропоморфизм в отношении технологии не нов; Но социальная робототехника, целенаправленно разрабатывающая антропоморфных роботов, похоже, усиливает эффект.

Антропоморфизация роботов вызывает самые разные опасения, и некоторые из них были идентифицированы, признаны и изучены в сообществах социальной робототехники и HRI. сами себя. Например, Теркл [16] и Шойц [17] выразили обеспокоенность по поводу эмоциональных отношений, которые люди развивают с такими роботами. Существует также знаменитая проблема « зловещей долины » . С тех пор как японский робототехник Масахиро Мори [18] представил (фрейдистскую) концепцию « жуткого » в робототехнике, возникла дискуссия о том, как роботы могут вызывать жуткое чувство и как этого можно избежать, например, при создании андроидов [19]. Для Фрейда сверхъестественное — это психологическое переживание странности в знакомом. В своем эссе 1919 года, посвященном этой концепции [20], он интерпретирует рассказ ЭТА Хоффмана, в котором фигурирует реалистичная кукла и другие « жуткие » элементы: люди чувствуют себя странно, когда объект, который должен был быть безжизненным, кажется живым и им угрожает их собственная опасность. Бессознательные, скрытые импульсы. В контексте робототехники « сверхъестественное » означает, что, когда робот выглядит похожим на человека, но движется, например, странными и пугающими способами, это может вызвать пользователей. Можно попытаться дать этому психологическое объяснение — по Фрейду или другое — например, страх смерти. А с антропологической точки зрения феномен антропоморфизации можно концептуализировать как форму технологического анимизма: анимизм, приписывание души или духа неодушевленным объектам и природным существам, тогда не понимается как нечто, принадлежащее « примитивной » стадии. нашей социальной эволюции или ограничивается « естественными » явлениями, а как нечто, что лежит в основе построения агентности и личности [21], и это также происходит по отношению к технологическим объектам, таким как роботы. Однако здесь я остановлюсь на антропоморфизме и конкретно на антропоморфизации.

Учитывая проблемы, возникающие в связи с антропоморфизмом, некоторые робототехники выступают за то, чтобы не ставить целью создание антропоморфных роботов как таковых. Гетц и его коллеги, например, утверждали, что роботов следует проектировать так, чтобы они соответствовали их задачам, а не считать человеческое подобие главной целью. [22] Тем не менее, поскольку создание антропоморфных роботов облегчает взаимодействие человека и робота и помогает достичь поставленных целей, например, в здравоохранении и уходе за пожилыми людьми, его часто преследуют разработчики социальных роботов.

В этой статье задается вопрос: как можно ответить на проблему антропоморфизма с точки зрения философии технологии? Существуют интересные возможности подойти к этой теме, некоторые из которых могут быть связаны с упомянутыми дискуссиями в области социальной робототехники. В этой статье я подхожу к этой проблеме, задавая вопрос, как философски сформулировать

отношения между роботами и людьми. Сначала я выдвинул две противоположные философские точки зрения по этому вопросу, которые я назову «наивным инструментализмом» и «некритическим постгуманизмом». В своей формулировке они могут выглядеть как подставные взгляды, но, как я покажу, есть исследователи, которые придерживаются, по крайней мере, более мягких версий. Я покажу, как каждая из этих точек зрения может быть связана с различными нормативными позициями по антропоморфизму и антропоморфизации: позиции, которые касаются не описания и интерпретации этих явлений, а вопроса, следует ли нам антропоморфизировать роботов и создавать антропоморфных роботов. набросок этих крайних противоположностей помогает мне разработать третий, альтернативный взгляд, подпитываемый герменевтическим, реляционным и критическим подходом.

Хотя я знаю, что это обсуждение также связано с В связи с темой робототехники и дискуссиями о моральном статусе роботов я не буду здесь напрямую затрагивать эти конкретные нормативные вопросы, а вместо этого сосредоточусь на теме антропоморфизма и связанных с ней проблемах отношений человека и технологии, имеющих отношение к социальной робототехнике и правам человека. Это не означает, что не существует никаких нормативно значимых выводов для социальной робототехники, которые можно было бы сделать из обрисованных позиций, что я и сделаю, но основное внимание в этой статье уделяется не моральному статусу. В конце статьи я рассмотрю последствия этого подхода для социальной робототехники и образования ученых-компьютерщиков и инженеров.

Хотя в статье говорится о работе в области социальной робототехники и прав человека, учитывая основную аудиторию этого журнала, стоит особо отметить, что эта статья представляет собой философское обсуждение различных взглядов на отношения человека и робота и нормативных позиций в отношении антропоморфизма и антропоморфизма; это не обзор литературы и не эмпирическое исследование.

Тем не менее, я надеюсь, что статья может стимулировать плодотворный диалог между социальными робототехниками и философами. Более того, обсуждение взглядов на взаимоотношения между роботами и людьми также актуально для размышлений об отношениях между технологиями и людьми в целом и, таким образом, может привлечь философов техники.

2. Два противоположных проблемных взгляда: Наивный инструментализм и некритичность Постгуманизм

Существует по крайней мере два противоположных взгляда на отношения между роботами и людьми, каждый из которых может быть связан с различными нормативными реакциями на антропоморфизм и антропоморфизацию:

2.1 Наивный инструментализм

Один из них — настаивать на том, что роботы — это машины и всего лишь инструкции для человеческих целей. Согласно этой точке зрения, которую я

Назвать «наивным инструментализмом» антропоморфизацию роботов — это своего рода психологическая предвзятость. Мы (учёные) знаем, что робот — это всего лишь инструмент, но тем не менее, когда мы взаимодействуем с роботом, наша психология (психология пользователей) заставляет нас воспринимать робота как некую личность. Пользователи думают, что робот похож на человека, хотя на самом деле это не так. (Мы, учёные, знаем это) Роботы — это всего лишь машины. Философская основа этой точки зрения двояка. Во-первых, он предполагает дуалистический взгляд на мир с точки зрения реальности и видимости (взгляд, который иногда приписывают Платону): предполагается, что в действительности робот — это инструмент, тогда как внешне он больше похож на человека. Наивный инструментализм также предполагает дуалистический взгляд на людей и нелюдей: людей и

Нелюди — это взаимоисключающие категории, и между ними существует глубокая онтологическая пропасть. Если рассуждать о технологии, это означает, что технологию можно найти на стороне нелюдей (например, вещей, объектов), тогда как люди совершенно другие (личности, субъекты). Роботы и люди относятся к совершенно разным онтологическим категориям.

Во-вторых, он предполагает версию метафизического и эпистемологического реализма: объекты существуют независимо от наших понятий и восприятий, и мы можем описать их объективным образом. Пользователей вводит в заблуждение внешний вид: на самом деле робот — это всего лишь машина. Ученые могут описать эту реальность объективно и изучить, как наше сознание вводит в заблуждение антропоморфические особенности робота, форма взаимодействия, а также наше социальное и культурное происхождение.

Эта точка зрения может привести по крайней мере к двум нормативным позициям в отношении антропоморфизма и задачи социальной робототехники. Во-первых, исследователи в области социальной робототехники могут использовать эту предвзятость, создавая человекоподобных роботов, чтобы улучшить взаимодействие человека и робота и лучше достичь своих целей. Не только робот является инструментом; Антропоморфизация сама по себе также является инструментом, который можно использовать. Именно эту позицию я обрисовал в начале этой статьи.

Оно предполагает неравномерное распределение знаний о реальном положении дел: разработчик робота знает реальность, в то время как пользователь — по крайней мере временно — введен в заблуждение, полагая, что робот — это человек. Можно сказать, что конструктор роботов работает как своего рода волшебник [23], создающий иллюзию того, что робот — человек. Как заметил философ дизайна Флюссер [24], сами термины «дизайн» и «машина» этимологически связаны с хитростью и обманом. Роботы созданы для того, чтобы производить антропоморфизацию пользователей, которых обманом заставляют поверить, что робот — это человек, или (принимая во внимание нынешний уровень развития робототехники), по крайней мере, готовы временно приостановить неверие и принять иллюзию. Согласно инструменталистской точке зрения, эта магия и хитрость антропоморфизации не проблематичны, если цель достигнута: если антропоморфизм робота и антропоморфизация пользователей приводят к лучшему взаимодействию человека и робота и если это, в свою очередь, достигает целей, которые хотели люди. достижения (например, конкретного здоровья

задача по уходу), то такой трюк разрешен и даже рекомендован. Это снова инструментализм в действии.

Обратите внимание, что по тем же соображениям, если цели не достигнуто антропоморфизмом, инструменталист не будет его использовать. Если робот по какой-либо причине не выполняет свою функцию, то инструменталистские рассуждения будут советовать не делать этого.

Другая нормативная позиция состоит в том, чтобы с помощью инструментализма утверждать, что, независимо от функционирования и эффективности робота, весьма проблематично антропоморфизировать машины и разрабатывать такие машины вообще и в принципе, поэтому считается и утверждается, что роботы — это нечто большее. инструменты, инструменты. Поэтому, как утверждается, специалисты по социальной робототехнике должны прекратить их разработку или, по крайней мере, убедиться, что пользователи осознают, что они являются всего лишь инструментами. Рассмотрим, например, позицию Брайсона: роботы — это инструменты, которые мы используем для достижения наших собственных целей [25]. Говорить о них таким образом, чтобы предположить, что они люди, вводит в заблуждение.

Более того, могут быть независимые причины, по которым антропоморфизм проблематичен. Например, можно утверждать, что создание таких машин также нежелательно из-за потенциальных этических проблем, упомянутых ранее: люди могут эмоционально привязываться к машинам и, как утверждают Спарроу и Спарроу, такие роботы могут оторвать их от реальности [26].] Например, социальный робот в сфере здравоохранения может ввести пользователей в заблуждение, заставив их думать, что он действительно жив или действительно может быть другом. Хотя нынешний уровень развития техники не создает этого для всех пользователей и в течение длительного времени, разработки в области социальной робототехники, похоже, направлены на эту иллюзию, например, с такими роботами, как робот-детеныш тюленя Паро, который используется в больницах и учреждениях ухода за больными. дома. Тёркл [16] утверждал, что такие роботы, как Паро, создают лишь иллюзию взаимоотношений.

С инструменталистской точки зрения можно было бы поддержать такую критику и добавить, что социальным робототехникам следует скорее

разрабатывать роботов, которые избегают этой иллюзии и работают просто как инструменты. Не следует создавать роботов, которые « притворяются » чем-то большим, чем просто инструментом. В случае реализации это вполне может стать концом самого проекта социальной робототехники в изложенном выше смысле, поскольку, если у робота больше не будет каких-либо особенностей, вызывающих антропоморфизацию (или если пользователю будут постоянно напоминать, что он представляет собой просто машина), неясно, как она может создать « естественное » социальное взаимодействие с людьми — или вообще « социальное » взаимодействие. С этой точки зрения нам следует прекратить разработку и использование социальных роботов, которые способствуют антропоморфизации. Если это означает, что взаимодействие человека

« естественный » или менее « социальный », тогда пусть будет так.

Но какую бы нормативную позицию в отношении антропоморфизации ни занимали (независимо от того, следует ли нам антропоморфизироваться или нет), обе позиции предполагают инструменталистский взгляд на отношения между роботами и людьми. Они оба думают, что роботы — это всего лишь инструменты; Они расходятся лишь в том, что касается оценки антропоморфизма и антропоморфизации. Более того, называть этот инструментализм « наивным » ни в коем случае нельзя.

Относится к незнанию антропоморфизма, не говоря уже о робототехнике. Обе нормативные позиции очень хорошо осознают, что такое антропоморфизм и что он делает (его природа и его последствия). Термин « наивный » в « наивном инструментализме » относится только к незнанию неинструментального измерения технологии. Что это значит, станет ясно в следующих разделах, особенно когда я распакую третью позицию.

2.2 Некритический постгуманизм

Другая точка зрения, находящаяся на другом конце спектра и условно названная « некритическим постгуманизмом », заключается в том, чтобы полностью принять социальных роботов как квазиличностей и « других ». Позвольте мне распаковать это. Постгуманизм может иметь несколько значений, но здесь я имею в виду теорию, которая критикует традиционные гуманистические мировоззрения, которые ставят человека в центр мира (антропоцентризм) и вместо этого расширяют круг онтологических и моральных интересов нечеловеческих существ. Это могут быть нечеловеческие животные, а также, например (некоторые) роботы. В отличие от инструменталистской позиции, здесь социальные роботы приветствуются как часть постгуманистической экологии или сети людей и нелюдей. Вместо дуалистического мировоззрения, противопоставляющего людей вещам, согласно этому взгляду, люди и нелюди являются частью одной и той же сети или экологии и запутаны различными способами.

Постгуманисты побуждают нас пересекать границы и искать дуализм и бинарность. Это включает в себя человеко-технологические

Бинарная нология. Согласно этой точке зрения, люди — это технологические существа, а технологии — это люди: люди всегда использовали технологии, люди создают технологии, а технологии являются частью нашего мира не только в инструментальном смысле. Например, мы можем говорить о технологии так, как будто это вещь, мы можем ожидать от технологии большего, чем она может предложить (что часто происходит в робототехнике, а также, например, в искусственном интеллекте), мы можем поклоняться технологии и т.д. Говоря о роботах, это означает, что, согласно этой точке зрения, роботов не следует рассматривать как дополнительные инструменты.

и вещи, если это означает, что единственное отношение, которое мы имеем к ним, является инструментальным; Вместо этого мы можем признать их частью нашего человеческого и культурного мира и даже « встретить » их как других: не как других людей, но и как социальных субъектов.

связи, сущности, которые разделяют наш мир. Более того, вместо реалистического взгляда на мир и на то, как мы познаем мир, постгуманисты склонны к нереализму: они верят, что у нас не может быть объективного взгляда на реальность (как если бы мы могли познать реальность, независимую от наших взглядов). способы познания) и что научные представления (например, о роботах) являются социальной конструкцией. Таким образом, то, что « есть » роботы, не исчерпывающе описано наукой и не может быть познано независимо от человеческой субъективности и культуры. Возьмем, к примеру, сами термины « робот » и « машина » : значение этих терминов недоопределено научными и инженерными определениями, поскольку они имеют свою собственную культурную историю.

и зависят от современного использования. Что означает слово « робот » , может зависеть, например, от социального контекста и соответствующего использования языка. Например, мы обычно не называем беспилотные автомобили роботами (в большинстве случаев мы называем их автомобилями), хотя согласно большинству технических определений они, безусловно, таковыми являются.

Подобный постгуманистический и социальный конструктивистский подход мог быть вдохновлен Харауэй, которая в своем « Манифесте киборгов» [27] и последующих работах выступала за пересечение границ путем включения машин и животных в политическую сферу, или Латура, чей немодный подход Современный подход включает нелюдей в социальное [28].

Точка зрения Харауэя касается не только создания буквальных киборгов (слияния людей и машин); Она стремится изменить наше представление о (важности) границ между людьми и машинами как метафизически, так и политически. Например, можно сказать, что роботы должны быть включены в общество. Аналогичным образом, по мнению Латура, роботы рассматриваются как нелюди, которых необходимо включить в

общественно-политический коллектив; Исключить их означало бы сохранить строгую границу между природой и культурой, а также между людьми и нечеловеческими существами. Несовременный подход выходит за рамки таких бинарных систем. Или можно интерпретировать исследования Гункеля, которые используют Левинаса для изучения того, можно ли считать роботов « Другими» [29, 30] как не тождественные (его точка зрения более критична), но потенциально ведущие к такой точке зрения. В традиционной феноменологии другой (иногда пишется: « Другой») противостоит самости и обычно относится к другим людям. Философ Эммануэль Левинас добавил к этому специфическое использование этого термина: « Другой» — это радикальный аналог « Я » , которое является абсолютно другим, и мы призваны этически реагировать на этого Другого. И все же Левинас все еще беспокоился о людях.

существа. С постгуманистической точки зрения тогда можно задаться вопросом: может ли этот другой или Другой также быть нечеловеческим животным или робот?

Согласно постгуманистической точке зрения, проект социальной робототехники не обязательно проблематичен и даже может быть принят, поскольку нам предлагается включить нелюдей в сферу социального, и подобные антропоморфизации могут в этом помочь. Нам следует выйти за рамки антропоцентризма инструменталистской позиции: человек не должен быть центром. Социальная робототехника может помочь в проекте децентрализации человека. Эта позиция также поддерживает антропоморфизм социальных роботов: это нормально и, возможно, даже желательно создавать роботов, которые выглядят как люди (антропоморфизм) и создаются и воспринимаются как таковые (антропоморфизация) не потому, что они являются инструментами для наших целей, а потому, что такие роботы предложить нам « другого» , с которым мы можем взаимодействовать и которого мы можем включить в соц. Постгуманизм отвергает антропоцентризм: он отвергает идею о том, что человек должен быть центром мира и моральной заботы. Человек

Таким образом, цели — не единственные, которые имеют значение, и мы также можем захотеть создать социальных роботов, которые вообще не являются антропоморфами.

Последняя позиция не обязательно отвергает антропоморфизм как таковой; оно просто указывает на то, что существует множество других возможностей. Однако, основываясь на взглядах Гункеля [29, 30] и придерживаясь постмодернистского подхода, можно было бы занять и другую, менее благоприятную нормативную позицию по отношению к антропоморфизации: можно утверждать, что, если смотреть с постгуманистического подхода, ориентированного на различие или инаковость, антропоморфизация В этом смысле восприятие роботов как людей проблематично в той мере, в какой оно не уважает отличие или инаковость робота, понимаемого как искусственный другой. Вместо того, чтобы проецировать нашу человечность на роботов (какие мы являемся людьми и какими мы видим себя людьми) и, таким образом, (зло)использовать их как простые проекционные экраны нашего образа себя, мы должны относиться к ним как к другим людям, имеющими собственные права и уважайте их разницу. Если мы не способны этого сделать, нам не следует строить антропоморфных роботов (но мы можем строить других). Таким образом, эта нормативная позиция также связывает антропоморфизм с антропоцентризмом, но отвергает оба.

2.3 Проблемы

И наивный инструментализм, и некритический постгуманизм проблематичны. Первый вариант наивен, если он не в состоянии полностью понять, что роботы — это не просто инструменты, но также и непреднамеренные действия. Последовательностей и связаны с людьми посредством опыта, языка, социальных отношений, повествований и т. д. Его дуалистический и реалистический взгляд на мир, который рассматривает отношения между людьми и технологией как внешние, не позволяет нам увидеть, что роботы — это не просто вещи (которые являются результатом технического конструирования и письма), но и в то же время человеческие существа. и социально сконструированный. Второй некритичен, если он забывает, что роботы — это машины, созданные человеком, которые вполне могут противостоять нам как « квази-другие» [31] , но никогда не являются полностью другими или полностью нечеловеческими, поскольку они создаются, интерпретируются и придают смысл людям. . Хотя они могут противостоять нам как внешний другой, они никогда не являются для нас полностью внешними. Фокус постгуманистов на инаковости и социокультурном конструировании роботов ведет к игнорировать их происхождение из человеческих и материальных практик. Роботы тоже машины. И эти машины сделаны

людьми. Другими словами, обе точки зрения не в достаточной степени признают, что социальные роботы внутренне и (как я скажу об этом ниже) внутренне связаны с людьми различными способами.

Я выражаю эту критику с точки зрения отсутствия полного понимания и забвения , поскольку оба вполне могут иметь некоторое знание о непредвиденных последствиях и происхождении технологии или в принципе могут приобрести эти знания. Точнее, поэтому: дело не в том, что наивные инструменталисты в принципе не знают, что их технология может иметь непредвиденные последствия или что

Некритические постгуманисты теоретически не знают, что технология создается. Они знают, что существуют эти отношения, или могут приобрести эти знания. Скорее, проблема в том, что они обычно не понимают более глубокого значения этих знаний для (мышления) отношений между людьми и роботами и, следовательно, для проблемы антропоморфизма.

В частности, обе точки зрения не понимают, насколько сильна на самом деле переплетенность людей и роботов, т. е. что это внутренняя связь (см. ниже).

Однако существует альтернативная, « третья » позиция, которая решает эту проблему, но при этом позволяет нам критически относиться к антропоморфизму и антропоморфизации, а также к проекту социальной робототехники. Как мы можем прийти к такому мнению?

Один из способов — попытаться найти « среднюю » позицию между двумя крайностями: социальные роботы не являются ни полностью инструментальными, ни совершенно другими. Но не ясно, что это значит. Каков « неинструментальный » аспект робота?

И что значит, что робот может быть в какой-то степени « другим » ? Поэтому, чтобы найти реальную альтернативу, я предлагаю подвергнуть сомнению и изменить предположения, поддерживающие обе позиции. В частности, я предлагаю отказаться от дуалистического предположения о том, что отношения между людьми и роботами являются внешними, и показать, как люди, социальные роботы и социальные

робототехника есть. Это создает новую, совершенно другую позицию, которая не является средней позицией, а выходит за пределы поля, определенного противоположными взглядами, изложенными здесь.

Глава 3. К третьему взгляду: критическому, реляционному и герменевтическому

Изложим такую позицию, которая применима к роботам вообще, но сформулирована в ответ на противоположные позиции, возникшие в ответ на антропоморфизм и антропоморфизацию в социальной робототехнике. Чтобы представлять собой реальную альтернативу, определенную выше, она должна состоять как минимум из следующих элементов:

3.1 Роботы — это люди, но не люди: их создают и используют, и они определяют наши цели

Во-первых, социальные роботы не просто появляются из дикой местности или из ниоткуда (как будто они являются частью « нечеловеческой » природы или как будто они появляются как пришельцы из космоса), но спроектированы и созданы людьми. Это подразумевает, что « инструмент » или « другой », выдвигаемый в двух рассмотренных нами позициях, не только нечеловеческий, но и человеческий одновременно. Роботы не только связаны с нашими целями и намерениями, как если бы они были внешними вещами, которые, так сказать, просто противостоят нам извне; они тоже созданы нами в конкретных научно-технических практиках и ими мы пользуемся . В этом смысле они люди. А это также означает, что они никогда не смогут быть полностью « другими » . А также есть

одинаковость. Даже роботы, не являющиеся антропоморфными, имеют человеческий аспект, поскольку мы их создали и используем. Их не нужно приветствовать и привлекать в « нашу » сферу; Они уже являются его частью с тех пор и по мере того, как мы их используем. И поскольку мы их создали, они никогда не являются просто инструментами, а наполнены нашими целями, нашими значениями и нашими ценностями. Это инструменты, но они наши инструменты. И своими непредвиденными последствиями они также формируют наши цели. Таким образом, инструмент и цель взаимозависимы.

Например, роботу можно придать антропоморфные черты, чтобы он мог функционировать в конкретном контексте здравоохранения. Цель связана со здравоохранением, например, с обеспечением возможности общения с пациентами. Но благодаря тому, что роботы общаются с пациентами, содержание этой цели меняется: то, что означает общение с пациентами, меняется с одного значения (когда люди являются теми, кто общается как лица, осуществляющие уход) на другое (общение между лицом, осуществляющим уход, и пациентом — это нечто это также можно сделать с помощью машины).

Более того, если мы (как пациенты, как лица, осуществляющие уход) забудем, что роботы созданы людьми и что мы придаем им значение, это может привести к некритическому принятию роботов и закрытию возможностей для изменений. Робот и взаимодействие следует воспринимать не как данность, а как изменчивую вещь. Позвольте мне сказать больше об этом придании смысла во взаимодействии.

3.2. Лингвистическая и социальная конструкция роботов

Во-вторых, даже если мы возьмем робота как данность, то есть таким, каким он нам кажется и как мы взаимодействуем с ним как пользователи (а дизайнеры тоже являются своего рода пользователями, хотя и особой категорией пользователей), мы должны еще раз подвергнуть сомнению предположение о том, что между людьми и роботами существует внешняя связь. Это можно сделать, рассмотрев то, что я предлагаю называть эпистемологическим антропоцентризмом. Это означает, что то, что мы говорим о роботе, что мы знаем о роботе и как мы относимся к роботу, всегда будет зависеть от человеческой субъективности, создания значений, повествований, языка, метафор и т. д. Независимо от того, сможем ли мы отойти от антропоцентризма в моральном или метафизическом смысле, принятие критического философского подхода (вслед за Кантом, феноменологией и большей частью философии двадцатого века) означает признание того, что нам всегда приходится « проходить через человеческое » , когда мы говорить, думать и взаимодействовать с социальными роботами. То, что « есть » робот, не зависит от того, что мы о нем говорим и как мы реагируем на робота при взаимодействии. Человек не только материально создает роботов, но и (в процессе разработки, использования и взаимодействия) « конструирует » их посредством языка и социальных отношений, что необходимо предполагать, когда мы думаем об этих роботах и взаимодействуем с ними. Это можно выявить с помощью критических философских и социально-научных усилий, которые включают временную и историческую перспективу, а также роль языка и метафоры [32–34]. Рассмотрим роль языка: если я дам социальному роботу личное имя, то

Я уже сконструировал робота как компаньона, другого — как имеющего определенный пол и т. д. [35]. Робот не обязательно перестает функционировать как инструмент, и его можно рассматривать как инструмент в другое время или с другой точки зрения. Но название, наряду с другими факторами, такими как внешний вид робота, определяет значение робота.

Кроме того, это значение связано с социокультурной средой, в которую встроен робот. Давая ему определенное имя, пользователи также могут подключиться к целой культуре именования и гендерной принадлежности. Например, присвоение женского имени социальному роботу, служащему в домашнем хозяйстве (в ответ на особенности робота, побуждающего к такому социальному поведению), не только « превращает» этого робота в своего рода человека, « с которым» я буду взаимодействовать. у способами, принятыми в моем сообществе и обществе; оно также отражает, подтверждает и распространяет культуру, в которой женщины должны играть такую роль. В этом смысле социальные роботы не являются ни « просто машинами» , ни совершенно « другими» ; Они различными способами связаны с человеческим обществом и человеческой культурой и становятся возможными благодаря этому обществу и этой культуре. В этом смысле роботы не являются « внешними» по отношению к людям и их целям, но между людьми и роботами существуют внутренние отношения.

3.3 Реляционность и создание смысла: роботы встроены в культурное целое

В-третьих, и это связано с предыдущими пунктами: альтернативный взгляд должен быть полностью реляционным. « Относительный» следует понимать здесь в следующих смыслах. Альтернативная позитивность должна признать реляционность роботов не только встроенной в более крупные технические системы, но и в более крупные социотехнические системы: как и другие технологии, роботы « переплетены с социальными практиками и системами значений людей» [36:195] . Таким образом, они « связаны» с такими практиками и системами значений. Можно также сказать, что они связаны с социальными отношениями [32] или что они встроены в культурное целое.

В своей недавней работе я выразил эту социально-культурную принадлежность роботов, используя концепции Витгенштейна « языковые игры» и « форма жизни» : как и использование слов, использование таких вещей, как роботы, также встроено в (технологические) игры и форму жизнь [33, 37]. Я использую термин « форма жизни» не в биологическом смысле, а интерпретирую (используя Витгенштейна) его как культурно определенные игры и практики, составляющие целое или целые, которые придают смысл нашей деятельности и технологиям. Эти целые позволяют нам понимать роботов и взаимодействовать с ними, а роботы, в свою очередь, способствуют или являются частью нашей деятельности по созданию смысла. Мы не только рассказываем о технологиях; Технологии также активно способствуют созданию смысла. Эта паутина или горизонт значений, в свою очередь, позволяет нам придавать смысл конкретному опыту и вещам. Это можно также подчеркнуть, обратившись к термину « нарратив» .

Используя нарративную теорию Рикёра, Вессель Рейерс и я [38] доказали, что, как и текст, технологии составляют нарративы, объединяя персонажей и события в значимое целое. Например, конкретное взаимодействие человека и робота можно поместить в более широкий нарратив о построении дружеских отношений с роботами: это можно сделать, рассказывая истории о роботе, но дело в том, что сам робот также активно способствует формированию этого нарратива. . Это можно сделать, например, с помощью антропоморфной формы робота, который создает персонажей и события (например, двое друзей встречаются и спрашивают, как дела).

В более общем плане можно сказать, что технологии выполняют герменевтическую функцию, поскольку способствуют созданию смысла (термин « герменевтика» здесь используется в отношении значения и интерпретации). Это также происходит, когда мы взаимодействуем с роботами и посредством нашего взаимодействия с роботами. Они не только являются объектом нашего создания значений и интерпретаций (например, я интерпретирую робота как « простую машину»), но также формируют то, как мы понимаем мир.

Например, домашний робот, который взаимодействует с людьми как слуга, может означать, что социальный мир — это мир хозяев и слуг, а пользователь — хозяин.

Таким образом, роботы — это не просто инструменты, если это означает, что они будут стоять вне сферы человеческой культуры; Вместо этого они включаются в нашу герменевтическую деятельность пассивным образом (мы говорим о них и придаем им смысл) и активным образом (они способствуют формированию смысла наших историй).

В качестве базовой философской основы этот реляционный и герменевтический подход подразумевает, что онтологический язык « есть» (философские утверждения или вопросы о том, что такое что-то или кто-то, например, « Что такое робот?»), используемый реалистами и инструменталистами, становится проблематичным, особенно если это « есть» не понимается в реляционном смысле. Не существует ни « вещи» или « инструмента» сами по себе, ни « другого» , не связанного с человеком. Все технологии, созданные и используемые людьми, связаны с людьми. Не только так называемые « социальные» роботы, но и все роботы уже являются социальными и реляционными в этом смысле, даже без или до того, как у них есть функции, которые считаются « социальными» . Например, промышленный робот встроен в языковые и технологические игры, принадлежащие определенному социально-индустриальному контексту, в котором в капиталистических условиях человеческие рабочие были заменены машинами. Но в случае с социальными роботами социально-реляционный характер роботов становится еще более очевидным; Они, так сказать, выставляют напоказ нашу социальную среду и культуру. Например, социальный робот с поддержкой искусственного интеллекта и голосовым интерфейсом будет « разговаривать» с нами способами (и мы будем разговаривать с ним способами), которые связаны с тем, как « мы» (в нашей семье, в нашем сообществе, в нашей стране, наше языковое сообщество и т. д.) говорить. Например, если ему удастся быть юмористическим, то это возможно только на основе использования общей (человеческой) базы знаний, опыта и навыков. А если бы ему удалось пошутить с помощью машинного обучения, то

робот научился бы играть в нашу игру. Успех его деятельности полностью зависит от реляционного целого, от формы жизни, которая должна быть предположена, чтобы робот был забавным и, в более общем плане, имел смысл. Социальные робототехники могут заставить робота имитировать это создание смысла и создание значения (здесь я буду использовать смысл и значение как взаимозаменяемые), позволяя роботу учиться на материале, который он находит во всемирной паутине, где корпусный текст материализует часть смысла, который вокруг. В области исследований искусственного интеллекта уже проводятся исследования этических проблемных аспектов использования языковых корпусов из Интернета; например, в этих корпусах и в нашем обычном языке существует гендерная предвзятость, которую ИИ имеет тенденцию воспроизводить [39]. Когда технология искусственного интеллекта будет все чаще внедряться в роботов, таких проблем может стать больше. Но если рассматривать его на уровне взаимодействия человека и робота, которое по определению включает в себя людей и, следовательно, также связано с играми и более широкими смысловыми горизонтами, взаимодействие всегда уже является (в реальном времени) деятельностью по созданию смысла. Опять же, во многих отношениях робот — это больше, чем просто инструмент или инструмент: это элемент деятельности и игр, которые имеют смысл и зависят от более крупных целостностей. Более того, робот не является полностью « другим » в том смысле, что, поскольку он действительно социален, в его отношениях с человеческим обществом и человеческой культурой имеется много сходства. Он может добиться успеха, то есть успешно работать в качестве социального робота, только в том случае, если существует достаточное сходство с точки зрения создания основы для общего значения. Инопланетный « социальный » робот не будет рассматриваться нами, людьми, как социальный, поскольку у него не будет всех этих связей с человеческим смыслом.

3.4 Отсутствие герменевтического контроля

В-четвертых, однако, сказать, что эти действия по созданию материалов и созданию значений являются человеческими (это то, что делают люди), не означает, что они полностью находятся под нашим контролем. Мы не полностью контролируем смысл, в том числе и в « чисто » человеческих ситуациях (которые редко бывают « чистыми » , поскольку обычно...

если не всегда – опосредованно другими технологиями и вещами).

Также возникает смысл, когда мы взаимодействуем с роботами, происходят встречи и события. Иногда появляются значения, которых мы не ожидаем; Это непреднамеренное появление смысла. Например, мы можем смеяться или испытывать страх, когда робот вдруг начинает звучать или выглядеть очень по-человечески.

Как показывают исследования HRI, роботы часто удивляют людей. В этом смысле и в таких случаях в них есть инаковость —

Или, по крайней мере, к взаимодействию и к роботу-отношению.

Разработчики и дизайнеры таких роботов, а также пользователи могут пытаться контролировать смысл (например, дизайнер может попытаться избежать

« сверхъестественных » эффектов), но герменевтический контроль никогда не бывает абсолютным. Когда люди взаимодействуют с миром и друг с другом, всегда есть место для сюрпризов и значений, которых мы не ожидали. В этом смысле роботы никогда не станут больше машинами, если слово « машины » ассоциируется со значением слова « механический » и если это

интерпретируется как « предсказуемый » . Взаимодействие человека и робота, как и любое другое взаимодействие, может иметь непредвиденные и неожиданные последствия. Дело не только в том, что в HRI наблюдается неожиданное поведение, как это было документально подтверждено (например, [40]); Люди могут и будут также понимать взаимодействие в

неожиданные способы. Например, жест робота, который должен быть дружелюбным, может внезапно быть истолкован как враждебный.

Такие неожиданные интерпретации неизбежны и являются частью того, как мы существуем в мире. Тем не менее, такие возникающие значения также не могут быть концептуализированы как представляющие собой полную инаковость, поскольку появление смысла - например, непреднамеренных значений, как в примере - снова полностью зависит от человеческого смыслообразования и смыслового опыта. Мы всегда сможем сравнить то, что делает робот, с чем-то, что мы знаем из (остального) человеческого мира и из нашей собственной социальной среды, с чем-то таким же .

и это относится к самому себе. Оно никогда не бывает совершенно другим или чем-то еще.

3.5 Мощность

Наконец, социальная робототехника, понимаемая как взаимодействие и тщательная человеческая деятельность по созданию смысла, социальному использованию и строительству материалов, всегда переплетается с властью.

Социальные роботы в использовании и взаимодействии — это не просто инструменты или чисто « технические » действия, они имеют социальные и политические значения и эффекты, включая аспект власти. Например, язык « рабов » , используемый Брайсоном [25], глубоко проблематичен [41], как недавно признал автор.

А социальные роботы в контексте здравоохранения могут поднять вопросы о политике здравоохранения без роботов: как дела обстоят сейчас и как сейчас обращаются с людьми, и каково качество и оправданность такого способа действий и такого способа лечения людей? Каким интересам отвечает внедрение этих социальных роботов в данный конкретный контекст? Власть и технологии связаны. Наивный инструментализм упускает из виду этот момент, поскольку он имеет тенденцию разделять технологию и человека.

соц. Безопасно относить социальных роботов к категории « технологий » , он непреднамеренно упускает из виду « человеческие » вопросы, касающиеся общества и власти. (И иногда инструментальный тализм не является наивным и намеренно ослепляет нас в отношении социальных отношений и отношений власти.) Но социальные роботы также не являются « другими » или « нелюдями » , если это означает, что они просто партнеры в приятных нейтральных к власти отношениях. Постмодернистская культурная и материальная игра (направление, которого, я полагаю, Гункель избегает, но которое может быть следствием по крайней мере одной интерпретации Гункеля или постмодернизма в стиле Харауэя). Принимая всерьез политическую сторону Харауэя, а также вдохновляясь Марксом [42] и Фуко [43], мы должны указать на опасность того, что постчеловеческие фантазии о взаимодействии с роботами могут маскировать то, как проект социальной робототехники может быть встроен в и вносить свой вклад в нарративы о доминировании и гендерном неравенстве, капиталистических социально-экономических системах и неолиберальных нарративах, а также на микро- и мезо-властные игры на

уровень человеческих взаимодействий и организаций. Например, роботы могут использоваться для сбора данных, которые продаются третьим лицам, их использование может привести к сокращению рабочих мест, они могут способствовать распространению проблемных способов обращения с женщинами, их производство может включать эксплуатацию людей, а роботы могут использоваться для демонстрации власти крупной компании или государства. Во всяком случае, эти проблемы, связанные с властью, показывают реальность и возможность постгуманистической антиутопии, а не утопии. Социальная робототехника вполне может представлять роботов как « других » и « друзей » ; но за кулисами (и на самом деле не так уж хорошо спрятанными) могут скрываться манипуляции, эксплуатация и дисциплина человека. Можно также рассмотреть трудовые отношения, которые делают возможным появление социальных роботов. Прославление инаковости и инаковости, а также аргументы в пользу включения нелюдей в социальное могут упускать из виду более широкий социально-экономический контекст и скрывать тот факт, что, как утверждает Фуко [44], власть присутствует во всех видах отношений, включая отношения, которые мы **имеем** с другими людьми при разработке и использовании социальных роботов. Сосредоточение внимания на работе как на « другом » может отвлекать людей от использования роботов для извлечения данных у людей и воздействия на « тела и души, мысли, поведение и образ жизни других людей » [45:18] . Например, социальные роботы могут использоваться для манипулирования людьми с целью заставить их покупать определенные товары и для стимулирования другого поведения, что отвечает интересам стороны, которая не видна в самом взаимодействии человека с роботом. За антропоморфной маской робота, представленной как игривое приглашение к нейтральной власти, приятному и полезному взаимодействию с квази-другим, скрывается серьезное лицо человеческой власти и человеческих властных отношений. Отношения – это не просто развлечение, а игры – это не просто игра; Такой взгляд на социальную робототехнику (и на человека) также открывает ящик Пандоры с проблемами, связанными с властью. Социальная робототехника — это игры (в смысле Витта-Генштейна, см. выше), и некоторые из них действительно очень серьезные игры: силовые игры. Тот факт, что это справедливо и для всех видов других цифровых инструментов, не является поводом отводить взгляд от силового аспекта этих игр, когда дело касается социальной робототехники.

3.6 Что дает третий взгляд: роботы как взаимосвязанные инструменты

Парадоксально, но, раскрывая человеческое измерение робототехники в ее использовании и разработке, эта позиция позволяет нам критически относиться к антропоморфизированным роботам. Именно потому, что социальные роботы глубоко связаны с человеком и обществом, мы можем указать на проблемную (или хорошую) антропоморфизацию и объяснить, почему она проблематична (или хороша), указывая на большое целое, в которое встроен робот, поскольку пример конкретной игры о гендере. В то же время, и это снова несколько парадоксально, эта позиция избегает наивной инструменталистской позиции, серьезно воспринимая роль роботов как инструмента создания смысла и инструмента, который является частью и со-образует более широкий контекст: роботы.

являются частью того, что я назвал — по аналогии с контекстом — « контекстологией » [37: 263]: точно так же, как тексты являются частью большого контекста, который формирует их значение, (другие) технологии также являются частью более широкого человеческого значимая и социальная реляционная среда, в которую они встроены и которая формирует их значение. Роботы — это инструменты; но они являются инструментами в отношениях: они всегда связаны с людьми и социокультурными полями, в которых они действуют. Наивный инструментализм предполагает поверхностные внешние отношения между людьми и их инструментом. Подход, который я разработал, отвергает это предположение. Моя ~~деконструкция~~ дискуссия дебатов об антропоморфизме опиралась на критический подход к предполагаемым внешним отношениям между людьми и роботами. Люди и роботы запутаны во многих отношениях; они внутренне и глубоко связаны.

4 последствия для борьбы с феноменом антропоморфизации и для социальных Робототехника как проект и сфера деятельности

Что означает этот подход и « третий » взгляд на отношения между людьми и роботами для оценки антропоморфизма в социальной робототехнике? Какая нормативная позиция в отношении этих явлений и практик вытекает из предыдущей критической дискуссии?

4.1 Роботы — это не другие и не просто машины

Во-первых, на основе развиваемой здесь альтернативной точки зрения кажется, что нам придется отвергнуть не только две исходные точки зрения (наивный инструментализм и некритический постгуманизм), но и связанные с ними нормативные позиции. Я предлагаю, чтобы антропоморфизм и антропоморфизация в социальной робототехнике не осуждались как выход за рамки того, чем должна быть машина (просто машина), и не использовались (неправильно) как шанс спроецировать инаковость на машину. Во многих смыслах, изложенных выше, социальные роботы — даже антропоморфные — могут восприниматься как другие, но не являются Другими с большой буквы « О » , как если бы они таинственным образом находились вне человеческого мира. Социальные роботы созданы человеком и являются частью более масштабных усилий по созданию человеческого смысла. Поэтому они не могут быть построены как совершенно другие. Здесь много общего из-за связи между роботом и контекстом его развития, а также использования мошеннических технологий. Мы можем воспринимать их как других, но не следует забывать, что они созданы человеком¹.

¹ Я мог бы добавить, что они также не являются « другими » в смысле друзей или товарищей, поскольку все, что мы знаем о человеческой дружбе и товариществе, это то, что их ценность заключается, среди прочего, в том, что друг или товарищ не созданы нами, а , действительно, в своей инаковости. Это невозможно создать, хотя во взаимодействии может возникнуть некоторая степень инаковости. Однако я не буду здесь развивать этот момент.

Но социальные роботы не являются чем-то большим, чем просто машинами. Антропоморфизация — это одна из вещей, которые мы делаем как люди, создающие смысл: даже если роботы не созданы для того, чтобы поощрять антропоморфизацию, всегда будет происходить « антропоморфизация » в том смысле, что значение роботов, включая социальных роботов, всегда будет меняться. быть связанным с более широкими человеческими практиками создания значений и более широкими смысловыми горизонтами, « формой жизни » (если снова использовать концепцию Витгенштейна, представленный подход не позволяет нам принять решение за или против конкретной антропоморфизации со стороны разработчиков. дизайнеры или пользователи. Но это помогает нам понять, что любая дискуссия по конкретному делу должна будет вызывать значения, выходящие за рамки конкретного взаимодействия: значения, извлеченные из нашего общества и нашей культуры, со всеми их измерениями, включая измерение власти. Это подразумевает, что позиция сторонников антропоморфизма внутри инструментализма также несостоятельна, если и поскольку исследователи, придерживающиеся этой позиции, не желают учитывать эту более широкую реляционную, значение и нормативность выполняемой ими работы по антропоморфизации.

4.2 Переосмысление целей социальной робототехники

Более того, изложенный реляционный и герменевтический подход также следует применить к нашему пониманию социальной робототехники как области. Социальную робототехнику следует понимать как междисциплинарную и мультидисциплинарную область, которая занимается не только созданием совокупностей реальных артефактов и программного обеспечения, называемых « роботами » (инструменталистская позиция), или созданием « искусственных других » (постгуманистическая позиция), но и созданием « искусственных других » (постгуманистическая позиция). также всегда одновременно с созданием человеческих, социальных, культурных и политических смыслов (и само утверждение, что робот — это « другой » , является одним из этих смыслов, имеющим политическое значение и которое должно быть проблематизировано).

Таким образом, хотя я указал на то, насколько проблематичны конкретные взгляды и связанные с ними нормативные позиции, настоящая статья не приводит к конкретному нормативному выводу относительно того, должны ли робототехники создавать антропоморфных роботов. В лучшем случае из этого можно было бы извлечь предостерегающий совет: если вы проектируете таких роботов, то имейте в виду, что вы с большей вероятностью присоединитесь к этой более широкой деятельности по созданию социального значения со всеми ее мощными аспектами или, по крайней мере, присоединитесь к этой более широкой мошеннической технологии, чтобы большей степени. По этой причине некоторые могут прийти к выводу, что безопаснее не разрабатывать таких роботов. Другие будут упорствовать в создании этих роботов, но тогда у них будет ответственность делать это таким образом, чтобы осознавать и принимать во внимание социально-культурные значения и последствия. Это подводит меня к следующему пункту об ответственности робототехников.

4.3 Ответственность робототехников, преподавателей, и Пользователи-граждане

С нормативной точки зрения признание этого реляционного, герменевтического и политического измерения социальных роботов и робототехники в целом подразумевает, что робототехники, ученые-компьютерщики, инженеры, психологи и все другие, вовлеченные в эту область, также несут ответственность за значения, которые они генерируют, и последствия социальных отношений. они производят. Это следует рассматривать не как что-то второстепенное, а как часть основного бизнеса социальной робототехники. Например, и разработчик аппаратного обеспечения робота, и разработчик программного обеспечения для распознавания речи, которое будет использоваться в работе, должны понимать себя не только как создатели технологического продукта, имеющего определенные функции (например, возможность вести разговор с роботом). человек в домашнем контексте), но также и как социальные субъекты, которые — через робота — вмешиваются в социальное игровое поле, которое включает в себя не только технические задачи (например, предотвращение или уменьшение гендерных предубеждений). Ответственность социальных робототехников тогда означает ответственность за то, чтобы робот и взаимодействие человека с роботом достигали своих целей, а также за то, чтобы делать это таким образом, чтобы проявлять ответственность перед другими людьми: это означает реагировать на пользователей и другие заинтересованные стороны, которые могут быть на него влияют непредвиденные последствия дизайна, кода и т. д., включая потенциальное герменевтическое/семиотическое воздействие. Это означает реагировать на социальные проблемы наряду с техническими и выяснять, как изменения на технической стороне могут способствовать смягчению последствий того, что происходит на социальном уровне.

Определить точную сферу ответственности за потенциальное герменевтическое/семиотическое и социальное воздействие роботов непросто. Например, невозможно полностью предсказать будущее, и остается неясным, в какой степени человек несет ответственность за удивительные, неожиданные задуманные значения. Как уже говорилось, не существует такого понятия, как полный, абсолютный герменевтический контроль – ни в робототехнике, ни где-либо еще в человеческой жизни и культуре. Но следует максимально повысить уровень герменевтического контроля от предполагаемых значений к непреднамеренным и потенциальным (будущим) значениям. Можно было бы ожидать от исследователей и дизайнеров робототехники (и других сотрудников компаний и организаций, в которых они работают), что они, по крайней мере, попытаются представить и оценить такие предполагаемые и потенциальные воздействия. Исследователи других дисциплин, в том числе гуманитарных и социальных наук, могли бы помочь, разработав и передав для этого методологии, например, методы создания сценариев и методы этической оценки.

Учитывая связь с социальным уровнем, ответственность за социальную робототехнику включает в себя принятие ответственности за значения и последствия социальной робототехники с точки зрения власти. Технические исследователи в области робототехники, возможно, не считают, что они имеют большое отношение к власти, поскольку власть обычно понимается как связанная с отношениями между

политические власти и граждане. Однако понимание Фуко направляет нас к проблемам власти во всех местах и на всех уровнях социальной структуры. Лаборатория робототехники не исключена из этой всепроникающей власти. Вместе с другими, кто принимает решения о робототехнике и использует или использует роботов в различных контекстах, те, кто разрабатывает и проектирует роботов, обладают определенной властью формировать смысл практик, истории людей и цели, которые мы ставим перед собой как люди. Для разработки антропоморфных роботов эта сила, возможно, более значима, чем в случае с другими типами роботов, поскольку антропоморфизм поддерживает социальные аспекты роботов, включая власть. С этой силой приходит ответственность.

Власть не только негативна: она может ограничивать и угнетать, но также может и способствовать. Власть может принимать форму расширения прав и возможностей: людям можно дать власть делать что-то, она может увеличить их потенциал. В контексте социальной робототехники вопрос заключается в том, как дать техническим специалистам возможность нести ответственность за свои творения. К сожалению, инженерам и ученым традиционно не хватает полного набора инструментов для этого, поскольку основное внимание уделяется техническим знаниям и навыкам. Это необходимо изменить. В методологическом отношении необходим более широкий междисциплинарный, а также междисциплинарный и трансдисциплинарный подход. Это может означать, например, что наряду с языками информатики и инженерными инструментами также используются инструменты гуманитарных и социальных наук. Сам проект « социальной робототехники» должен быть задуман в реляционном ключе и связан с проектом критической оценки человеческой культуры и общества. Это необходимо, поскольку, как я предположил, лаборатория робототехники связана с более широкой социальной культурной среды, языка, власти и т. д., а создание и использование роботов в то же время формирует все это поле отношений в той же степени, в какой оно само формируется им. Если робототехники примут этот вызов и возьмут на себя ответственность и заново определят свое место в этом более широком целом, то они смогут сыграть важную роль в формировании будущего наших социальных связей — то, что всегда делали инженеры, от мостов и железных дорог до электричества и мобильной связи. телефоны — этически и социально ответственным образом.

Эти нормативные и методологические последствия подхода, изложенного выше, требуют трансдисциплинарности и реформы образования инженеров-робототехников и ученых-компьютерщиков. Когда они вступают в свою профессиональную жизнь, они должны быть не только готовы, но также готовы и способны связать свою практику, свои разработки и проектные решения с более широкой общественной дискуссией и культурным дискурсом, в которых они должны иметь возможность критически участвовать, используя слова и вещи (роботы). Но ответственность следует возлагать не только на робототехников. Граждане-пользователи, со своей стороны, также должны взять на себя ответственность за будущее социальной робототехники и, в более общем смысле, за будущее технологий и будущего общества. Им необходимо понять (и быть

обучены понимать), что социальные роботы, с которыми они могут столкнуться в некоторых контекстах и сопутствующих технологиях своей повседневной жизни, не являются ни « другими» , ни просто машинами или инструментами, а инструментами, которые являются частью их общества и культуры, ради которых они также нести ответственность. Если мы будем рассматривать роботов не как противоположность человеческой сфере, а как ее важнейшую часть, тогда роботы станут интересны так, как мы, возможно, не могли себе представить раньше. Объединение людей и языков из разных миров посредством трансдисциплинарного образования может привести к измененному пониманию того, чем технологии могут быть и чем они могут помочь обществу, и, в конечном итоге, к глубоким изменениям обоих. Именно к этому должны быть готовы студенты: студенты естественных и инженерных специальностей, а также студенты гуманитарных и социальных наук.

4.4 Общий вывод

В этом мета-смысле я полагаю, что нам нужно больше антропоморфизации, а не меньше. Не обязательно антропоморфизация из-за конструктивных особенностей роботов и возможностей взаимодействия, которые они обеспечивают, что усиливает и дополняет и без того неизбежные и существующие отношения между роботом и его социальной и культурной средой и, следовательно, может быть как захватывающим, так и очень проблематичным (для пример, потому что он способствует распространению определенных властных отношений, которые мы считаем нежелательными или неприемлемыми). Это требует изучения и обсуждения со стороны проектировщиков и пользователей-граждан, которые должны нести совместную ответственность за проект. Но нам особенно необходима большая антропоморфизация социальных роботов в смысле большего признания того самого понимания, что существует связь между роботами и их герменевом.

В такой среде существует глубокая социальная и семантическая взаимосвязь: понимание — как в лаборатории, так и в обществе в целом — что социальные роботы, как и любые другие технологии, действительно очень человечны. Помимо всего остального, антропоморфизм в социальной робототехнике демонстрирует эту реляционность и человечность: он напоминает нам, что технология всегда была человеческой.

Финансирование Финансирование открытого доступа предоставлено Венским университетом.

Декларации

Конфликт интересов Автор заявляет об отсутствии конфликта интересов.

Открытый доступ. Эта статья распространяется по международной лицензии Creative Commons Attribution 4.0, которая разрешает использование, совместное использование, адаптацию, распространение и воспроизведение на любом носителе и в любом формате при условии, что вы укажете соответствующую ссылку на оригинального автора(ов) и источник. предоставить ссылку на лицензию Creative Commons и указать, были ли внесены изменения. Изображения или другие материалы третьих сторон в этой статье включены в лицензию Creative Commons, если не указано иное.

в противном случае в кредитную линию к материалу. Если материал не включен в лицензию Creative Commons на статью и предполагаемое использование не разрешено законодательством или превышает разрешенное использование, вам необходимо будет получить разрешение непосредственно от правообладателя. Чтобы просмотреть копию этой лицензии, посетите <http://creativecommons.org/licenses/by/4.0/>.

Рекомендации

- Фонг Т., Нурбахш И., Даутенхан К. (2003) Обзор социально интерактивных роботов. Система автономности роботов 42: 143–166
- Даутенхан К. (2007) Социально интеллектуальные роботы: аспекты взаимодействия человека и робота. *Philos Trans R Soc Lond B Biol Sci* 362: 679–704
- Даффи Б.Р. (2003) Антропоморфизм и социальный робот. Система автономности роботов 42: 177–190
- Исигуро Х. (2006) Интерактивные гуманоиды и андроиды как идеальные интерфейсы для людей. В материалах 11-й международной конференции по интеллектуальным пользовательским интерфейсам (IUI '06). Ассоциация вычислительной техники, Нью-Йорк, штат Нью-Йорк, США, 2–9.
- Бартнек С., Куллк Д., Крофт Э. (2008). Инструменты измерения антропоморфизма, анимации, привлекательности, воспринимаемого интеллекта и воспринимаемой безопасности роботов. *Int J Soc Robot* 1: 71–81
- Злотовски Дж., Праудфут Д., Йогисваран К. и др. (2015) Антропоморфизм: возможности и проблемы во взаимодействии человека и робота. *Int J of Soc Robotics* 7: 347–360. <https://doi.org/10.1007/s12369-014-0267-6>
- Кислер С., Хиндс П. (2004) Введение в специальный выпуск о взаимодействии человека и робота. *Hum-Comput Interact* 19 (1): 1–8
- Ривз Б., Насс С. (1996) Уравнение СМИ: как люди относятся к компьютерам, телевидению и новым медиа как к реальным людям и местам. Издательство Кембриджского университета, Кембридж
- Кислер С., Пауэрс А., Фасселл С.Р., Торри С. (2008) Антропоморфные взаимодействия с роботом и роботоподобным агентом. *Soc Cogn* 26 (2): 169–181
- Гегель Ф., Гизельманн С., Петерс А., Холтхаус П., Вреде Б. (2011). К типологии значимых сигналов и сигналов в социальной робототехнике. В: Материалы международного семинара IEEE по интерактивному общению роботов и людей, стр. 72–78.
- Хухенбрандт Д., Эйссел Ф., Бобингер С., Нойфельд М. (2013) Когда членство робота в группе имеет значение. *Int J Soc Robot* 5 (3): 409–417
- Крегелох К.У., Бхаратарадж Дж., Кутти СКС, Ирмаля П.Р., Хуанг Л. (2019) Анкеты для измерения приемлемости социальных роботов: критический обзор. *Робототехника* 8(4):88. <https://doi.org/10.3390/robot840088>
- Пека А., Кокельберг М., Симут Р., Костеску С., Пинтеа С., Дэвид Д., Вандерборгт Б. (2016) Роботизированная терапия для детей с аутистическими расстройствами: измерение этической приемлемости. *IEEE Technol Soc Mag* 35(2):54–66
- Эйссел Ф. (2017) Экспериментальный психологический взгляд на социальную робототехнику. Автономная система робота 87: 363–371
- Дарлинг К. (2017) Кто такой Джонни? Антропологический подход во взаимодействии, интеграции и политике человека и робота. В: Лин П., Дженкинс Р., Эбни К. (ред.) Этика роботов 2.0. Издательство Оксфордского университета, Нью-Йорк
- Теркл С. (2011) Наедине вместе: почему мы ожидаем большего от технологий и меньше друг от друга. Основные книги, Нью-Йорк
- Шойц М. (2012) Опасности, присущие односторонним эмоциональным связям между людьми и социальными роботами. В: Лин П., Эбни К., Беки Дж. (ред.) Этика роботов. MIT Press, Кембридж, Массачусетс, стр. 205–222.
- Мори М. (1970) Зловещая долина. Энергия 7:33–35
- Макдорман К.Ф., Исигуро Х. (2006) Необъяснимое преимущество использования андроидов в исследованиях в области когнитивных и социальных наук. *Интерактивный изучение* 7: 297–337
- Фрейд С. (2003) Жуткое. Пингвин Классикс, Лондон
- Ричардсон К. (2018) Технологический анимизм. В: Суонкэт К., Мазард М. (ред.) Анимизм за пределами души: онтология, рефлексивность и создание антропологических знаний. Берган, Нью-Йорк, стр. 110–128.
- Гетц Дж., Кислер С., Пауэрс А. (2003) Сопоставление внешнего вида и поведения робота с задачами по улучшению сотрудничества человека и робота. В: Материалы 12-го международного семинара IEEE по интерактивному общению роботов и людей (ROMAN 2003), 2003, стр. 55–60.
- Кокельберг М. (2018) Как описать и оценить феномен «обмана»: переосмысление метафизики, этики и политики ИКТ с точки зрения магии и производительности и принятие реляционного и повествовательного поворота. *Этика Inf Technol* 20(2):71–85
- Флюссер В. (1999). Форма вещей: философия дизайна. Реак Книги ции, Лондон
- Брайсон Дж. (2010) Роботы должны быть рабами. В: Уилкс Ю. (ред.) Тесное взаимодействие с искусственными спутниками. Джон Бенджаминс, Амстердам, стр. 63–74.
- Воробей Р., Воробей Л. (2006) В руках машин? Будущее ухода за пожилыми людьми. *Умы Маха* 16: 141–161
- Харауэй Д. (2000) Манифест киборга. В: Дэвид Б., Кеннеди Б.М. (ред.) Читатель киберкультуры. Рутледж, Лондон, стр. 291–324.
- Латур Б. (1993) Мы никогда не были современными. Издательство Гарвардского университета, Кембридж, Массачусетс, Кэтрин Портер (транс)
- Gunkel DJ (2007) Думая иначе: философия, коммуникация. Издательство Университета Пердью, Вест-Лафайет, Технологии
- Gunkel DJ (2017) Другой вопрос: могут и должны ли роботы иметь права? *Этика ИнфоТехнол* 20:87–99
- Иде Д. (1990) Технологии и жизненный мир. Университет Индианы Пресс, Блумингтон
- Кокельберг М. (2012) Растущие моральные отношения: критика приписывания морального статуса. Пэлгрейв Макмиллан, Нью-Йорк
- Кокельберг М. (2017) Использование слов и вещей: язык и философия технологий. Рутледж, Нью-Йорк
- Кокельберг М. (2019) Движимые машинами: метафоры производительности и философия технологий. Рутледж, Нью-Йорк
- Кокельберг М. (2011) «Ты, робот: о лингвистическом построении искусственных других» в: *AI & Society* 26(1): 61–69.
- Джонсон Д.Г. (2006) Компьютерные системы: моральные сущности, но не моральные агенты. *Этика ИнфоТехнол* 8:195–204
- Кокельберг М. (2018) Технологические игры: использование Витгенштейна для понимания и оценки технологий. *Научно-английская этика* 245:1503–1519
- Кокельберг М., Рейерс В. (2016) Нарративные технологии: философское исследование нарративных возможностей технологий с использованием нарративной теории Рикера. *Хум Стад* 39: 325–346
- Калискан А., Брайсон Дж. Дж., Нарайанан А. (2017). Семантика автоматически извлекается из языковых корпусов, содержащих предубеждения, подобные человеческим. *Наука* 356 (6334): 183–186.
- Леменьян С., Финк Дж., Мондада Ф., Дилленбург П. (2015) Вы делаете это неправильно! изучение неожиданного поведения при взаимодействии ребенка с роботом. В Тапус А., Андре Э., Мартин Дж.К., Ферланд Ф., Амми М. (ред.) Социальная робототехника. ICSR 2015. Конспекты лекций по информатике, том 9388. Спрингер, Чам.
- Кокельберг М. (2015) Трагедия мастера: автоматизация, уязвимость и дистанция. *Этика ИнфоТехнол* 173:219–229
- Маркс К. (1867/1976) Капитал (Том I). Penguin Books, Лондон (переиздание 1990 г.).
- Фуко М. (1980) Власть/знание: избранные интервью и другие произведения 1972–1977. Колин Гордон изд. Книги Пантеона, Нью-Йорк

44. Фуко М (1980) Власть/знание. Гордон С. (редактор) Гордон С., Маршалл Л. (транс). Книги Пантеона, Нью-Йорк
45. Фуко М. (1988) Технологии самости: семинар с Мишелем Фуко. Мартин Л.Х., Гутман Х., Хаттон П.Х. (ред.). Издательство Массачусетского университета, Амхерст

Примечание издателя Springer Nature остается нейтральной в отношении заявлений о юрисдикции в опубликованных картах и институциональной принадлежности.