

■ Концептуальная исследовательская статья / Conceptual article

# Искусственный интеллект как антропологический проект. От редукции к становлению (ч. 2)<sup>1</sup>

Д. Е. Николаев  
АНО Исследовательский центр «Аналитик». Екатеринбург, Россия

**АННОТАЦИЯ** Современные исследователи активно обсуждают, чем искусственный интеллект (ИИ) является для человека: культурным орудием, средством цифровой экстернализации, цифровым значимым Другим или альтернативой человека. Гораздо реже встречаются размышления о том, кем является сам человек для технологий и их создателей. В работе показано, что технологии ИИ, представляющие собой не единое инженерное решение, а конгломерат различных подходов, не только базируются на принципиально разных образах человека, но и транслируют их через собственную архитектуру, поддерживая развитие соответствующих практик. Тем самым данные решения влияют на культуру, конструируют новые миры и нового человека в них. Архитектура большинства внедряемых и обсуждаемых исследователями типов ИИ (генеративный, человекоцентричный, «заслуживающий доверия», объяснимый, интерпретируемый, воплощенный, ситуативный, энактивистский, социальный, а также проектируемый в рамках социотехнического и акторно-сетевого подходов) опирается на редуцированные образы человека. Одним из ключевых рисков, связанных с широким внедрением подобных систем, является опасность расчеловечивания – «вымывания» из человеческих практик контекстуального понимания, смыслотворчества, моральной и нравственной ответственности, глубоких межличностных связей, со-действия, вариативности, автономии и самодетерминации. В результате человек становится не более чем исполнителем социальной роли (пользователь, клиент, работник, пациент, студент и т. п.). Преодоление редукции человека в системах ИИ возможно через изменение дизайна архитектуры. Соответствующие технологии должны выходить за пределы упрощенных антропологий и строиться на сложных образах человека. Такие технологии способны не только фасилитировать становление человека, развитие его потенциала, но и способствовать формированию по-настоящему человекоцентричной культуры – культуры достоинства (А. Г. Асмолов). В работе сформулированы практические ориентиры, которые должны лечь в основу проектирования таких архитектур.

**КЛЮЧЕВЫЕ СЛОВА** искусственный интеллект, потенциал человека, человекоцентричность, образы человека, антропология технологий

**ССЫЛКА ДЛЯ ЦИТИРОВАНИЯ** Николаев, Д. Е. (2026). Искусственный интеллект как антропологический проект: от редукции к становлению (ч. 2). *Образовательная политика*, 24(1), 96–108.

© Николаев Д. Е., 2026  
Контент доступен под лицензией  
Creative Commons Attribution 4.0 License.  
The content is available under  
a Creative Commons Attribution 4.0 License.

Поступила: 20.11.2025  
Принята: 01.12.2025  
Дата публикации: 05.04.2026



**Дмитрий Евгеньевич НИКОЛАЕВ**  
к. псих. н., заместитель директора,  
АНО Исследовательский центр «Аналитик»  
(620219, РФ, Екатеринбург, пр. Ленина, 38а),  
<https://orcid.org/0000-0002-3385-1640>,  
e-mail: dnpsy2019@gmail.com

<sup>1</sup> Начало см. в журнале «Образовательная политика» 23(4), 2025.

*Воплощенный, ситуативный  
и энактивистский ИИ (Embodied / Situated /  
Enactive AI)*

Противовесом использованию ИИ как «черного ящика знаков» выступают традиции, подчеркивающие роль тела, среды и практики в становлении интеллекта. Работы по воплощенному и ситуативному познанию показывают, что ментальное во многом является результатом непрерывных взаимодействий субъекта с физической и социальной средой (Suchman, 1987; Brooks, 1991).

Идеи о том, что когнитивные процессы неотделимы от тела и среды, сегодня реализуются в робототехнике и системах человеко-машинного взаимодействия. В робототехнической практике наблюдается активный прогресс: автономные роботы-помощники, системы для логистики, промышленные и медицинские роботы проходят этапы коммерциализации, хотя их массовое использование остается ограниченным. Научные лаборатории исследуют подход энактивизма, но их воплощение в прикладных сервисах тормозится сложностью сенсорной интеграции, вопросами безопасности поведения роботов в реальном мире и высокими затратами на испытания в полевых условиях. Социокультурные барьеры (недоверие, регуляторные ограничения) также сдерживают массовое внедрение.

Архитектура воплощенного ИИ должна интегрировать сенсорно-моторные слои, локальную обработку восприятия и непрерывные петли обратной связи. Интеллект здесь рассматривается как эффект взаимодействия тела и среды, а не только как

решение задач. Такой дизайн подчеркивает, что когнитивные способности возникают на практике, а не только в теории (Brooks, 1991; Suchman, 1987; Pfeifer & Bongard, 2006).

Традиции воплощенного и ситуативного познания рисуют образ человека как существа, чье представление о мире и себе самом возникают в активности тела, в локальных практиках и в конкретных средах взаимодействия. Исходя из этого, ИИ рассматривается не как источник символов, а как компонент поведенческой взаимосвязи. Машины и интерфейсы становятся средой,

в которой формируются умения, навыки и смысл. В этой парадигме успешность технологии измеряется по тому, насколько она способствует развитию практических компетенций, а не по степени автоматизации действий (Brooks, 1991; Suchman, 1987). Образ человека здесь достаточно сложен, но часто упрощается в конкретных инженерных практиках, где, как правило, доминирует

символическая репрезентация.

## НАРЯДУ С ИНСТИТУЦИОНАЛЬНЫМИ И ОНТОЛОГИЧЕСКИМИ КАТЕГОРИЯМИ АКТИВНО РАЗВИВАЕТСЯ ОБЛАСТЬ ИССЛЕДОВАНИЙ, РАССМАТРИВАЮЩАЯ ОТНОШЕНИЯ «ЧЕЛОВЕК — МАШИНА» КАК ДОЛГОВРЕМЕННЫЕ, ЭМОЦИОНАЛЬНО И ЭТИЧЕСКИ НАСЫЩЕННЫЕ ПРАКТИКИ

*Социотехнический и акторно-сетевой подход  
к ИИ (Sociotechnical / Actor-Network approach)*

Акторно-сетевая теория и смежные подходы переводят внимание с локального субъекта на сети отношений. ИИ в этой перспективе выступает как актант в сети людей, артефактов, стандартов, рынков и институтов. Такой подход позволяет показать, как технические артефакты участвуют в формировании траекторий распределенной власти, и как бренд, инфраструктура и нормативные практики становятся частью единого социотехнического поля (Latour, 2005; Law, 1992). Методологически это предполагает картирование связей, этнографию разработчиков, модераторов

и пользователей, а также рассмотрение материальных условий — от дата-центров до правовых режимов.

Архитектура ИИ в этом случае должна включать материальную инфраструктуру, процессы сбора данных, каналы распространения и правила использования. Это означает проектирование не только алгоритмов, но и сетей взаимодействия между элементами системы (поставщики данных, регуляторы, пользователи, сервисные организации и т. п.), с учетом того, что ИИ сам становится актантом в социальной сети. Ключевая задача для разработчиков — проектировать механизмы распределенной ответственности, прозрачности и институционального контроля по всей сети.

Поскольку такой подход переводит фокус с индивидуальной природы на структурную позицию человека в сети взаимодействий с артефактами, институтами и инфраструктурами, человек понимается здесь как участник гетерогенной сети интересов и действий. Такой образ полезен для анализа институциональных эффектов, но он полностью игнорирует психологическое в человеке, не говоря уже о процессах самоактуализации и становления личности.

#### *Социальный ИИ (Social / Relational AI)*

Наряду с институциональными и онтологическими категориями активно развивается область исследований, рассматривающая отношения «человек — машина» как долговременные, эмоционально и этически насыщенные практики. Исследования в области affective computing демонстрируют, что ИИ потенциально может сканировать, интерпретировать и симулировать эмоциональные состояния, что открывает поле для новых форм взаимодействия (Picard, 1997). Основное внимание здесь сосредоточено не на способности ИИ решать те или иные задачи, а на отношениях и долговременных социальных связях между человеком и ИИ.

На теоретическом уровне эти подходы обладают устойчивой академической базой. За последнее десятилетие появилось множество работ, исследующих, как

алгоритмы модифицируют практики модераторов контента, образовательные процессы, клинические консультации и административные процедуры (Gillespie, 2018; Eubanks, 2018; Crawford, 2021). В академическом сообществе наблюдается активный рост исследований в данной сфере, но их влияние на инженерные практики остается ограниченным из-за институциональной разобщенности между инженерами, дизайнерами и исследователями (Lam et al., 2024; Prasad, 2024<sup>2</sup>; NTIA, 2024<sup>3</sup>).

Развитие чат- и компаньонных ботов, а также терапевтических и социально-ориентированных агентов (Woebot, Wysa и пр.) служит важным примером того, как ИИ становится частью повседневной эмоциональной и социально-психологической среды. Для части задач (психологическая поддержка, первичная терапия, мотивация) накоплена клиническая и эмпирическая база, включающая рандомизированные исследования и реальные внедрения, в то же время академические дискуссии подчеркивают ограничения, связанные с долгосрочными эффектами, рисками зависимости и этикой доверия (Eltahawy et al., 2024; Iglesias et al., 2023; Foran et al., 2024). В результате компаньонные ИИ находятся на стадии активной коммерциализации и клинической апробации, но с существенными ограничениями по уровню качества и по регулируемым стандартам.

Архитектура социального ИИ изначально ориентирована на долгосрочное поддержание взаимоотношений, поэтому, помимо модулей синтеза и распознавания речи, она должна включать модели репрезентации диалоговой истории, механизмы адаптации к индивидуальным стилям и протоколы поддержания эмпатических стратегий. Интерфейс и поведение системы ориентированы на создание и поддержание социального контекста, а не только на однократную полезность. В такой архитектуре

2 Prasad, V. (2024, August 2). *AI algorithm audits: Key control considerations*. ISACA Industry Guidance.

3 NTIA. (2024, March). *AI accountability policy report*. National Telecommunications and Information Administration.

значимы этические гарантии (ограничение манипуляций, прозрачность роли ИИ) и процедуры клинической и социальной валидации (Bickmore & Picard, 2005; Picard, 1997).

Исследования в области affective computing и социальных роботов предполагают образ человека как существа, чья идентичность и морально-нравственное развитие формируются в отношениях, насыщенных эмоциональными обменами. Технологии, которые симулируют или реагируют на эмоции, апеллируют к этому образу, предлагая поддержку, заботу или терапевтическую помощь. Это порождает вопросы о доверии, заботе, моральной значимости «симулированной эмпатии» и рисках эмоциональной эксплуатации. В широко распространенных LLM это проявляется в феномене «социально-эмпатического» диалога, когда модели способны имитировать отклик, вызывающий у людей чувство понимания, хотя при этом реальная эмпатия у системы, естественно, отсутствует (Picard, 1997; Bender et al., 2021).

## Редукция человека в системах ИИ

Большинство уже реализованных и пока лишь обсуждаемых в научной и философской литературе типов ИИ, в том числе рассмотренные выше, опираются на редуцированные, неполные образы человека. Редукция образов обусловлена, конечно, в какой-то мере необходимостью формализации, экономической логикой и научно-технологическими ограничениями, но еще и инертностью мышления, стереотипами, склонностью к опрощению жизни (Асмолов и др., 2016) и оперированием бытовыми представлениями о человеке.

Такие образы функционально удобны для машинного обучения и внедрения. Они позволяют задать входы, цели и функции оптимизации, но одновременно они исключают из рассмотрения ключевые измерения человеческого (Bommasani et al., 2021; Bender et al., 2021) – например, идентичность, временную перспективу и деятельностный характер развития личности, преадаптивность, диалогический и даже полифонический характер сознания, мышления и идентичности, со-действие, самодетерминацию, способность

к саморазвитию, групповую, культурную вариативность и рефлексию, то есть те слои человеческого опыта, которые трудно оцифровать.

Даже когда исследователи и инженеры обсуждают «человекоцентричный ИИ», исходя из их представления о человеке, можно говорить не о подлинной человекоцентричности,

а о частных, редуцированных «центричностях», поскольку в центре внимания находится не человек во всей его сложности, а те или иные отдельные его проявления или аспекты: потребности, здоровье, безопасность, благополучие, способность решать задачи и внедрять инновации и т. п. (Николаев, 2025).

Рассмотрим ключевые варианты редукции образов человека в технологиях ИИ.

### *Редукция через формализацию: экономия на сложности*

Техническая практика проектирования ИИ сопряжена с задачей формализации, т. е. перевода сложных человеческих феноменов в набор входных данных, метрик и целевых функций. Именно этот перевод порождает первую форму редукции. Человек в инженерной парадигме предстает как совокупность наблюдаемых признаков

**БОЛЬШИНСТВО УЖЕ  
РЕАЛИЗОВАННЫХ И ПОКА ЛИШЬ  
ОБСУЖДАЕМЫХ В НАУЧНОЙ И  
ФИЛОСОФСКОЙ ЛИТЕРАТУРЕ ТИПОВ  
ИИ, В ТОМ ЧИСЛЕ РАССМОТРЕННЫЕ  
ВЫШЕ, ОПИРАЮТСЯ НА  
РЕДУЦИРОВАННЫЕ, НЕПОЛНЫЕ  
ОБРАЗЫ ЧЕЛОВЕКА**

и предикторов, полезных для задачи: пользовательская цель, поведенческие предпочтения, метрики удержания, клики, ответы на тесты и т. п. Такая операционализация необходима для обучения моделей и оптимизации, но при этом она исключает из поля зрения аспекты, не поддающиеся количественному измерению. Эта логика лежит в основе подходов, где человеческое становится совокупностью предпочтений или задач, которые нужно учитывать или в которые нужно встроиться (Russell, 2019). Верифицируемость и управляемость становятся главными ценностями, системы оптимизируются под краткосрочные KPI и экономические интересы.

*Редукция через функционализм:  
человек как когнитивный ресурс*

В ряде подходов (особенно тех, что ориентированы на автоматизацию и оптимизацию рабочих процессов) человек редуцируется до роли функционального, производственного ресурса — источника данных, поставщика труда, потребителя услуг. Человеческая жизнь рассматривается через призму ее пригодности для генерации прибыли предприятия (Zuboff, 2019; Crawford, 2021). Здесь редукция приобретает специфический характер: человек и его деятельность становятся не более чем «новой нефтью» для обогащения. Подобные подходы игнорируют ценности, смыслы, субъектность, потенциал и рассматривают индивида преимущественно как капитал, ресурс, объект рыночной оценки (Асмолов, 2023а; Николаев, 2025).

*Редукция через герменевтику  
«без понимания»: дискурсивные модели  
и иллюзия смысла*

Генеративные модели и LLM породили новую форму редукции: человек как понимающий и смыслообразующий субъект подменяется человеком — потребителем выданных алгоритмом смысловых фрагментов. LLM воспроизводят вероятностные паттерны корпуса текстов, создавая иллюзию понимания и осмысленности, но уязвимость этой стратегии заключается в отсутствии

подлинной рефлексии и трансляции смыслов, формирующих идентичность человека (Bommasani et al., 2021; Bender et al., 2021). Следствием такой «герменевтической» редукции становится ситуация, когда люди начинают воспринимать сгенерированные машиной тексты как легитимные источники смысла и знания, что может подрывать критическое мышление и самостоятельную интерпретацию. Здесь же возникает и социальное расслоение как непосредственный эффект редукции, когда голоса и переживания определенных групп систематически маргинализируются в текстах (Fricker, 2007).

*Редукция через телесность:  
ограниченность воплощенных подходов*

Подходы, подчеркивающие роль воплощенного и ситуативного познания, справедливо критикуют редукцию к символам и показывают, что ментальное зачастую возникает в практиках тела и в конкретных средах (Brooks, 1991; Suchman, 1987). Тем не менее и в них существует редукция, хотя и иной природы, заключающаяся в игнорировании вопросов нарративной интеграции, целостной личности и морального развития. Воплощенный субъект может превосходно действовать в контексте, но оставаться неспособным содействовать формированию автономной идентичности человека, чьи ценности и смыслы строятся во времени и в культуре. Иными словами, способность «делать» не равнозначна способности «становиться».

*Редукция через распределение: риск  
технологической экстернализации личности*

Теория расширенного сознания предлагает плодотворную антиредукционистскую перспективу, утверждая, что внешние артефакты могут становиться частью широкой когнитивной системы (Clark & Chalmers, 1998). Однако эта перспектива привносит в рассмотрение человека собственную редукцию, предполагающую, что личность и ее развитие становятся объектами внешней поддержки, что снижает значимость самого человека. Если технологии превращаются

в непреодолимый когнитивный «аутсорсинг», становление личности может быть вынесено вовне и низведено до механики взаимодействия с интерфейсом.

#### *Редукция через эмоции: симуляция против формирования*

Технологии эмоционального взаимодействия и социальные роботы (боты) формируют образ человека как эмоционального субъекта, сигналы которого можно считывать, чтобы реагировать на них. Однако симуляция эмпатии алгоритмическими системами не обязательно приводит к развитию эмпатии у человека и не заменяет в полной мере межличностное общение. Кроме того, постоянное использование эмоциональных «подпорок» может препятствовать развитию способности переносить фрустрацию, регулировать эмоции и вступать в подлинные субъект-субъектные отношения, критически важные для становления личности.

Редукция проявляется здесь в упрощении эмоциональной сложности человека до набора сигналов, доступных для программной реакции (Picard, 1997).

В целом можно говорить, что система ИИ опирается на редуцированный образ человека, если ее успех определяется преимущественно краткосрочными метриками (удержание, клики, скорость); если она оперирует строго формализуемыми целями без механизма обновления ценностных ориентиров пользователя; если она создает когнитивную или эмоциональную зависимость; если в ней отсутствуют процедуры лонгитюдной оценки влияния на личность, а также если экономическая модель предусматривает эксплуатацию труда человека в ущерб его праву на развитие.

## Механизмы, через которые редукция фасилитирует процессы расчеловечивания

Редуцированные образы человека в проектных практиках ИИ приводят к множеству пересекающихся последствий. С педагогической и психологической стороны это создает риск «выращивания» зависимых субъектов, лишенных опыта преодоления, что мешает развитию автономии и морально-нравственной зрелости. С социальной точки

зрения это усиливает несправедливость: те, чьи голоса и способы существования плохо репрезентированы в корпусе данных, на которых обучался ИИ, оказываются маргинализированными в системах, претендующих на универсальность знаний (Fricke, 2007; Benjamin, 2019). Кроме того, генеративные модели создают иллюзию авторитета и могут подрывать способность к

## ГЕНЕРАТИВНЫЕ МОДЕЛИ И LLM ПОРОДИЛИ НОВУЮ ФОРМУ РЕДУКЦИИ: ЧЕЛОВЕК КАК ПОНИМАЮЩИЙ И СМЫСЛООБРАЗУЮЩИЙ СУБЪЕКТ ПОДМЕНЯЕТСЯ ЧЕЛОВЕКОМ — ПОТРЕБИТЕЛЕМ ВЫДАННЫХ АЛГОРИТМОМ СМЫСЛОВЫХ ФРАГМЕНТОВ

критическому мышлению, транслируя предубеждения, отраженные в корпусе текстов (Bommasani et al., 2021; Bender et al., 2021).

Перечисленные и неперечисленные эффекты редукции можно обобщить, назвав их «расчеловечиванием» (Dang et al., 2025; Schultz et al., 2024). Использование редуцированных образов человека во все более широко взаимодействующих с ним системах ИИ подрывает возможность индивидуального проектирования человеком себя как личности и как создателя смысла. Это проявляется в том, что личностный выбор подменяется «оптимизированными» решениями, а человеческий поступок оказывается лишь откликом на алгоритмически сгенерированные стимулы. Эту ситуацию можно назвать отчуждением и утратой

подлинности (Сартр, 2000) или опытом абсурда (Камю, 1990). Таким образом, технологическая редукция порождает не просто вытеснение человека и утрату рабочих мест или навыков, но и более фундаментальную угрозу утраты человеческого в человеке, прежде всего, способности конституировать смысл своей жизни.

Редуцированные, неполные образы человека приносят практическую пользу для решения узких задач, но одновременно искажают представление о том, что значит быть человеком, и тем самым способствуют социальным, антропологическим и психологическим рискам. Мы рискуем создать мир, где технологические артефакты не просто обслуживают существующие практики, но постепенно формируют людей под критерии эффективности, измеримости и коммерческой воспроизводимости, то есть создают человека определенного типа, зачастую неспособного быть рефлексивным и автономным субъектом. Такой транслируемый технологией образ человека, как и многие образы в истории, словно «джин, выпущенный из бутылки, может перестать слушаться хозяина, подняться на уровень мировоззрения и официальной идеологии, стать нормативным канонem, а затем начать вести свою работу, творя людей по своему подобию» (Асмолов, 2023б).

Перечислим основные механизмы расчеловечивания при использовании редуцированного образа человека.

*Эпистемическая редукция и утрата контекста.* Первое следствие редукции — это сужение эпистемического поля. Когда решения принимаются на основе статистики, без исторического и нарративного контекста, то реальные мотивации, смыслы и тонкие различия человеческих практик теряются. Это проявляется в медицинских, образовательных и социальных алгоритмах, где автоматические прогнозы заменяют понимание. Это ведет к тому, что человеческая сложность оказывается опрошенной (Асмолов и др., 2016; Eubanks, 2018; Bommasani et al., 2021). Это прямо способствует игнорированию индивидуальности и вариативности, что может приводить к систематической

маргинализации тех, чьи жизненные траектории не укладываются в усредненные шаблоны (Fricker, 2007).

*Инструментализация и экономизация человеческой деятельности.* Редуцированные образы человека тесно сочетаются с экономическими подходами, где человек рассматривается как ресурс, источник данных и инструмент извлечения прибыли (Zuboff, 2019; Crawford, 2021). Когда человек в технологии воспринимается всего лишь как поставщик данных, «юзер» или «трафик», то технологии ориентируются на поддержку тех форм поведения, которые удобны рынку, а вовсе не на повышение качества жизни и становление личности.

*Деградация навыков и моральная «дескиллизация».* Делегирование задач машине, от простых рутинных навыков до сложных практических и этических решений, может со временем ослаблять способность человека выполнять те же действия без технической поддержки и лишать его навыков нравственной оценки собственной деятельности (Vallor, 2016). Если технологии систематически «избавляют» человека от необходимости принимать трудные решения, это может вести к утрате контекстов, в которых формируется и воспитывается моральная зрелость (Kohlberg, 1982; Фромм, 1990; Арендт, 2008).

*Нормализация и стандартизация поведения.* Этот механизм связан с тем, что технологии, оптимизирующие определенные метрики, создают «нормы» поведения. Как показывают исследователи, *когда метрика становится целью, она перестает быть хорошей метрикой*, поскольку люди забывают о реальных целях и перестраивают свою деятельность под метрику, искажая социальные процессы, особенности которых она призвана измерять. Эта закономерность хорошо известна в науке, была описана различными исследователями и носит разные наименования: закон Гудхарта, закон Кэмпбелла и т. п. (Strathern, 1997; Campbell, 1979). В случае с использованием систем ИИ происходит следующее: поведение, ценное в рамках конкретной модели (например, максимизация кликов или соответствие предписанным профилям),

начинает вытеснять и деформировать формы взаимодействия, которые ранее потенциально могли бы служить гуманитарным или образовательным функциям.

*Имитация эмпатии и редукция межличностных отношений.* Развитие систем, симулирующих эмпатию и социальную поддержку, создает новую форму «отношений», в которой субъекты получают эмоциональные подкрепления от алгоритмических «партнеров». Это может быть в определенной степени полезно (максимальная скорость предоставления и доступность услуг), но одновременно

существует риск подмены межличностной заботы автоматизированными взаимодействиями, что уменьшает потребность в сложных человеческих навыках взаимодействия и поддержания социальных связей (Turkle, 2011; Picard, 1997). Если взаимодействие с машиной

систематически замещает человеческие контакты, это меняет социальную ткань и может подрывать способность общества формировать глубокие межсубъектные отношения.

*Политические и институциональные эффекты: власть и ответственность.* Редукция человека облегчает институциональное усреднение ответственности: когда решение считается «автоматизированным» или «экспертным», общество склонно доверять системе или перекладывать ответственность на производителей, не предоставляя субъектам реальных средств оспаривания решений. Это институционально упрощает власть, уменьшает пространство для социального участия и сужает формы социального, коллективного, а также создает возможности для злоупотребления властью.

Бурное развитие моделей ИИ, редуцирующих представление о человеке, способствует расчеловечиванию особенно, когда

технологии проектируются и внедряются без внимания к контексту, без механизмов защиты приватности, без верификации целей, а также когда экономическая логика диктует приоритеты краткосрочной монетизации. В таких условиях перечисленные выше механизмы особенно способствуют исключению из человеческих практик тех сторон, которые делают их по-настоящему человеческими.

## Преодоление редукционизма

### ПОВСЕМЕСТНОЕ РАСПРОСТРАНЕНИЕ ТЕХНОЛОГИЙ, ПОСТРОЕННЫХ НА РЕДУЦИРОВАННОМ ОБРАЗЕ ЧЕЛОВЕКА И ТРАНСЛИРУЮЩИХ ЕГО, ЗАСТАВЛЯЕТ ЗАДУМАТЬСЯ: А НУЖНО ЛИ В ПРИНЦИПЕ ИСКАТЬ ЧТО-ТО ИНОЕ?

Повсеместное распространение технологий, построенных на редуцированном образе человека и транслирующих его, заставляет задуматься: а нужно ли в принципе искать что-то иное? Стоит ли усложнять? Может быть, редукция здесь вполне адекватна и уместна?

Исследователи отмечают, что начавшийся во второй половине XX в. антропологический поворот в философии и науке разрушил прежний образ человека, который строился путем формулирования той или иной его природы. Вместо этого возникла задача нового самоопределения, свободного от редукции к какому-либо началу (Смирнов, 2017). Изменяются не только базовые ориентиры прежде разрозненного знания о человеке, но и подчеркивается отказ от поиска конечной природы его существования. Новым предметом трансдисциплинарных исследований в науках о человеке становится сложный, изменяющийся человек в изменяющемся мире (Асмолов и др., 2022).

Исследователи показали, что в бизнесе редукционизм проявляется в моделях мотивации, ориентированных на вознаграждения и показатели продуктивности, когда работники становятся «ресурсом», «средством»,

«юнитами производительности», что ведет к эмоциональному выгоранию, снижению производительности, вовлеченности и инновационности, а также к росту текучести кадров (Асмолов, 2024; Николаев, 2024). В образовании алгоритмы адаптивного тестирования, фокусирующиеся на краткосрочных результатах, могут подавлять развитие критического мышления, творчества и нравственной рефлексии. В социальной сфере автоматизированные системы часто репрезентируют уязвимые категории людей как «риски», а не как субъектов со своими историями и потенциалами (Eubanks, 2018). В медицине редукция человека до тела закрепляет практики, унижающие человеческое достоинство, игнорирующие важность социальных связей.

Более того, само обращение к таким мейнстримным понятиям современной науки и практики, как потенциал человека, вариативность, долгосрочная ориентация, жизнестойкость, вовлеченность и потоковые состояния, а также человекоцентричность, важнейшим показателем которой является индивидуальное пространство выбора или широта возможностей каждого (Асмолов и др., 2024; Николаев, 2025), возможно только с опорой на сложностное, нередуцированное представление о человеке (Sen, 1999; Nussbaum, 2011; Асмолов, 2023а; Леонтьев и др., 2011; Николаев, 2025). Это понимают не только академические исследователи, но и люди исключительно практического склада мышления, например представители бизнеса. Один из наиболее известных консультантов по организационному управлению С. Кови писал по этому поводу: «Наличие парадигмы целого человека как инструмента описания человеческой природы предоставляет вам редкую возможность объяснить, предсказать и диагностировать серьезные проблемы в вашей жизни и в вашем подразделении или группе. Если у лидеров имеются неточные и неполные парадигмы человеческой природы, они создают системы (например, системы коммуникации, набора кадров, премирования, оплаты работы, подготовки), которые не в состоянии задействовать потенциал людей полностью» (Кови, 2008).

В этой связи архитектуры ИИ, спроектированные на основе нередукционистских образов человека (например, понимаемого как исторически, социально и культурно обусловленный субъект становления), открывают совершенно новую проектную перспективу, поскольку они направлены не на замену человека машиной и не только (и не столько) на развитие его способностей с помощью внешнего культурного орудия, а на сопровождение становления личности, на расширение его потенциала, а также на поддержку автономии и субъектности. В различных нередукционистских традициях становление человека объясняется разными, но не взаимоисключающими механизмами: свободный выбор и проект самого себя (экзистенциализм), герменевтический диалог с культурными, историческими и языковыми традициями (герменевтика), взаимодействие в сетях дискурсивных практик и властных отношений (постструктурализм), самоорганизация в сложных системах (теории сложности). Современные синтетические подходы в определенной мере интегрируют эти представления и расширяют их до проблемы социальных возможностей. Так, теория *sarabilities* нобелевского лауреата по экономике А. Сена и философа М. Нуссбаум связывает человеческое развитие с реальным доступом к возможностям жить достойной жизнью, что задает нормативную меру для оценки технологий (Sen, 1999; Nussbaum, 2011).

Эти теоретические подходы не только позволяют диагностировать угрозу расчеловечивания в результате редукции. Также они указывают понятный путь преодоления редукционизма: признание свободы, ответственность за выбор и проектирование условий, в которых субъект может развиваться. Архитектура ИИ, ориентированная на образ «человека становящегося», не только способствует развитию потенциала человека и формированию человекоцентричной культуры, она, содействуя развитию человеческого в человеке, способна «снять» большую часть страхов и опасений, связанных с развитием ИИ. Если технология перестает

использоваться и восприниматься как конкурент человеку, если ее цель — помощь в его становлении, то дальнейшее ее развитие приводит не к ненужности человека как вида, не к появлению киборгов или технологической сингулярности, а к развитию человеческого в нас. Это в корне трансформирует проблему: развитие технологий перестает быть угрозой для человеческой сущности и становится инструментом ее усиления. Такой подход предлагает практическую схему минимизации и перевода страхов в проектные задачи: как проектировать системы, институты, культуру и экономику таким образом, чтобы технологический прогресс способствовал человеческому развитию. Внедрение подобных архитектур не гарантирует полного отсутствия проблем, но они делают наиболее пугающие сценарии значительно менее вероятными, переводя развитие ИИ из области отчуждения в область совместного человеческого становления.

Из вышеизложенных философских и антропологических подходов вытекают проектные принципы таких систем. Во-первых, это многоуровневость целей: архитектура должна комбинировать краткосрочные функции (помощь в решении конкретной задачи) и долгосрочные практики становления (поддержка и фасилитация автономии, субъектности, самодетерминации, развитие компетентности, формирование нарративной целостности, уникальность траекторий становления). Во-вторых, это опора на представления о зонах ближайшего и вариативного развития. Поддержка должна быть не только временной и вариативной, но и своевременно сворачиваемой, чтобы избежать закрепления зависимости и поощрять самостоятельность (Выготский, 1991; Wood et al., 1976). В-третьих, необходимы архитектурные и организационные принципы, которые гарантируют человеку контроль над информацией и историей ее использования (суверенитет данных и provenance) (Floridi, 2013). В-четвертых, архитектура должна предусматривать встроенную объяснимость, при этом важен педагогический характер объяснений, которые должны иметь образовательную

функцию, помогать пользователю понимать, почему система дает тот или иной совет, и стимулировать критическое мышление (Miller, 2019; Gunning, 2019). В-пятых, должны присутствовать культурная чувствительность, гибкость и т. н. «партисипативный дизайн». Т.е. сообщество пользователей должно участвовать в формулировании разнообразия целей и критериев «успеха» становления, чтобы архитектура не транслировала унифицированных «единственно верных» представлений.

## Выводы

Проведенный анализ показывает, что обсуждение внедряемых технологий ИИ не может ограничиваться бинарной оппозицией замещения или развития человека. Принципиально важным является вопрос о том, какого именно человека предполагает и конструирует та или иная архитектура ИИ и какие практики она тем самым нормализует. Смещение исследовательского фокуса с показателей эффективности на реконструкцию антропологических предпосылок технологий задает методологическую оптику, в которой ИИ рассматривается не как конкурент человека, а как потенциальный участник совместного процесса становления.

Архитектуры систем искусственного интеллекта выступают носителями имплицитных теорий человека. Они «скриптят» поведение, формируют ожидания и тем самым участвуют в производстве новых форм субъектности. Архитектурные решения в ИИ следует рассматривать не только как технические, но и как социально-конструктивные: они влияют на то, какие формы практик и субъектности оказываются поддержанными, а какие вытесненными. В этом смысле проектирование ИИ является актом создания социальных форм и требует соответствующей исследовательской и нормативной ответственности.

Сравнительный анализ ключевых типов ИИ показывает, что даже подходы, декларирующие ориентацию на человека,

в большинстве случаев опираются на редуцированные, инструментальные его образы. Такие редукции функционально удобны и экономически эффективны, однако они систематически воспроизводят риски расчеловечивания практик, смещая акцент с человеческого становления на измеримость, управляемость и воспроизводимость. В долгосрочной перспективе это способствует формированию типов субъектности, плохо совместимых с автономией, рефлексией и смыслообразованием.

Поскольку архитектуры ИИ действительно кодируют антропологию, изменение этих архитектур следует рассматривать как один из ключевых способов влияния на траектории человеческого становления. Технологии, претендующие на служение человеку, должны выходить за пределы упрощенных образов человека и опираться на сложностное представление о нем как о субъекте становления. Такая ориентация предполагает не замену человеческой способности к выбору и смыслу, а создание условий, в которых эта способность может разворачиваться и укрепляться.

Тем самым задается альтернативная модель технологического развития, в которой эволюция искусственных систем может идти рука об руку с развитием человеческих способностей к смыслообразованию, ответственности, самодетерминации и саморазвитию. Реализация данной модели требует трансдисциплинарного синтеза и институционализации сотрудничества между инженерами, исследователями человека и сообществами пользователей на всех этапах жизненного цикла ИИ. Тем самым архитектура ИИ становится не просто объектом инженерного выбора, а инструментом целенаправленного, осмысленного участия в проекте человеческого становления.

## Список источников

Аверина, А. Ж., Александрова, Л. А., Васильев, И. А., Гордеева, Т. О., Гусев, А. И., Дергачева, О. Е., Иванченко, Г. В., Калитеевская, Е. Р., Курганская, М. В., Лебедева, А. А., Леонтьев, Д. А., Мандрикова, Е. Ю., Митина, О. В.,

Осин, Е. Н., Плотникова, А. В., Рассказова, Е. И., Фам, А. Х., & Шапкин, С. А. (2011). *Личностный потенциал. Структура и диагностика*. Смысл.

Арендт, Х. (2008). *Банальность зла. Эйхман в Иерусалиме*. Европа.

Асмолов, А. Г. (2023а). *Институт потенциала человека: сканирование горизонтов. Образовательная политика*, 2(94), 5–6. <https://elibrary.ru/gebgsa>

Асмолов, А. Г. (2023b). *Психология личности. Культурно-историческое понимание развития человека* (6-е изд.). Смысл.

Асмолов, А. Г. (2024). *Корпоративная культура: система подчинения или ключ к развитию каждого? Образовательная политика*, 19(4), 8–12. <https://elibrary.ru/sqzxcw>

Асмолов, А. Г., Шехтер, Е. Д., & Черноризов, А. М. (2016). *Что такое жизнь с точки зрения психологии: историко-эволюционный подход к психофизической проблеме. Вопросы психологии*, 2, 3–23. <https://elibrary.ru/wialdp>

Асмолов, А. Г., Шехтер, Е. Д., & Черноризов, А. М. (2022). Антропологический поворот: восхождение к сложности. В Л. П. Киященко, Т. А. Сидорова (Ред.), *Человек как открытая целостность* (с. 33–53). Академиздат. <https://elibrary.ru/frwque>

Асмолов, А. Г., Шехтер, Е. Д., & Черноризов, А. М. (2024). Совместная эволюция сообществ и их лидеров: восхождение к человекоцентричности. *Вопросы психологии*, 70(5), 3–15. <https://elibrary.ru/yzrsvb>

Выготский, Л. С. (1991). Динамика умственного развития школьника в связи с обучением. В *Педагогическая психология* (с. 391–410). Педагогика.

Камю, А. (1990). *Бунтующий человек. Философия. Политика. Искусство*. Политиздат.

Кови, С. (2008). *Восьмой навык: Руководство пользователя*. Альпина Бизнес Букс.

Николаев, Д. Е. (2024). *Четыре степени свободы: От мотивации персонала к вдохновлению соратников*. СберУниверситет, Академия потенциала Человека.

Николаев, Д. Е. (2025). Человекоцентричность и потенциал человека: экзистенциальный ракурс. *Новые психологические исследования*, 5(2), 30–50. <https://elibrary.ru/nerzqg>

Сартр, Ж. П. (2000). *Бытие и ничто: Опыт феноменологической онтологии*. Республика.

Смирнов, С. А. (2017). Антропологический поворот: его смысл и уроки. *Философия и культура*, 2, 23–35. <https://elibrary.ru/ygfhxh>

Фромм, Э. (1990). *Искусство любить. Исследование природы любви*. Педагогика.

## References

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21). Association for Computing Machinery, 610–623. <https://doi.org/10.1145/3442188.3445922>

- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press.
- Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2), 293–327. <https://doi.org/10.1145/1067860.1067867>
- Bommasani, R., et al. (2021). *On the opportunities and risks of foundation models*. arXiv preprint. <https://arxiv.org/abs/2108.07258>
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1–3), 139–159. [https://doi.org/10.1016/0004-3702\(91\)90053-M](https://doi.org/10.1016/0004-3702(91)90053-M)
- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67–90. [https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X)
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1093/analysis/58.1.7>
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Dang, J., & Liu, L. (2025). Dehumanization risks associated with artificial intelligence use. *American Psychologist*. Advance online publication. <https://doi.org/10.1037/amp0001329>
- Eltahawy, L., et al. (2024). Can robots do therapy?: Examining the efficacy of a CBT bot in comparison with other behavioral intervention technologies in alleviating mental health symptoms. *Computers in Human Behavior: Artificial Humans*, 2, Article 100035. <https://doi.org/10.1016/j.chbah.2024.100035>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Floridi, L. (2013). *The ethics of information*. Oxford University Press.
- Foran, H. M., et al. (2024). An automated conversational agent self-help program: Randomized controlled trial. *Journal of Medical Internet Research*, 26, Article e53829. <https://doi.org/10.2196/53829>
- Fricke, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 40(2), 44–58. <https://doi.org/10.1609/aimag.v40i2.2850>
- Iglesias, M., Sinha, C., Vempati, R., Grace, S. E., Roy, M., Chapman, W. C., & Rinaldiet, M. L. (2023). Evaluating a digital mental health intervention (Wysa) for workers' compensation claimants: Pilot feasibility study. *Journal of Occupational and Environmental Medicine*, 65(2), e93–e99. <https://doi.org/10.1097/JOM.0000000000002762>
- Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Indiana University Press.
- Kohlberg, L. (1982). Moral development. In J. M. Broughton & D. J. Freeman-Moir (Eds.), *The cognitive developmental psychology of James Mark Baldwin: Current theory and research in genetic epistemology* (pp. 277–325). Ablex Publishing Corp.
- Lam, K., Lange, B., Bili-Hamelin, B., Davidovic, J., Brown, S., & Hasan, A. (2024). A framework for assurance audits of algorithmic systems. *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*. arXiv preprint. <https://doi.org/10.48550/arXiv.2401.14908>
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems Practice*, 5, 379–393. <https://doi.org/10.1007/BF01059830>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Nussbaum, M. C. (2011). *Creating capabilities*. Harvard University Press.
- Pfeifer, R., & Bongard, J. C. (2006). *How the body shapes the way we think: A new view of intelligence*. MIT Press.
- Picard, R. W. (1997). *Affective computing*. MIT Press.
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.
- Schultz, M. D., Clegg, M., Hofstetter, R., & Seele, P. (2024). Algorithms and dehumanization: A definition and avoidance model. *SSRN Electronic Journal*, 40, 2191–2211. <http://dx.doi.org/10.2139/ssrn.5013103>
- Sen, A. (1999). *Development as freedom*. Oxford University Press.
- Strathern, M. (1997). 'Improving ratings': Audit in the British University system. *European Review*, 5(3), 305–321.
- Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.
- Verbeek, P.-P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

## Artificial Intelligence as an Anthropological Project. From Reduction to Becoming (Part 2)

Dmitry E. Nikolaev

PhD in Psychology., Deputy director of ANO Research Centre "Analytic" (38a, Lenin Ave., Yekaterinburg, Russian Federation, 620219), <https://orcid.org/0000-0002-3385-1640>, e-mail: [dnpsy2019@gmail.com](mailto:dnpsy2019@gmail.com)

**ABSTRACT** Scholars actively debate what artificial intelligence is for humans – a cultural tool, a means of digital externalization, a digital significant Other, or a future alternative, replacement for the human. Far less frequently discussed is who the human is for technologies and their designers. The paper shows that AI technologies – understood not as a single engineering solution but as a system of heterogeneous approaches – not only rest on fundamentally different conceptions of the human, but also transmit these images through their architectures, thereby facilitating corresponding practices and shaping culture, new worlds and a new forms of being. The architectures of most types of AI both, implemented in practice and only discussed in the literature (generative, human-centered, trustworthy, explainable, interpretable, embodied, situated, enactive, social, as well as designs developed within sociotechnical and actor-network frameworks) rely on reduced, incomplete images of the human being. Among the principal risks associated with broad deployment of such systems is the risk of dehumanization – the elimination from human practices of those features that make them genuinely human: contextual understanding, meaning-making, moral and ethical responsibility, deep interpersonal bonds, co-action, variability, autonomy and self-determination. In these architectures the human too often appears only as a performer of a social role (user, client, worker, patient, student, etc.). Overcoming the reduction of the human in AI architectures is possible through redesign of those architectures. Corresponding technologies must move beyond simplified anthropologies and be built on complex images of the human being. Such technologies can not only facilitate human development and the realization of human potential, but also contribute to the formation of a truly human-centered culture – a culture of dignity (A. G. Asmolov). Practical tenets intended to underlie the design of non-reductionist architectures are formulated.

**KEY WORDS** artificial intelligence; human potential; human-centricity; conceptions of the human; anthropology of technology

**TO CITE** Nikolaev, D. E. (2026). Artificial Intelligence as an Anthropological Project: From Reduction to Becoming (Part 2). *Educational Policy*, 24(1), 96–108.

© Nikolaev D. E., 2026

Received: 01.12.2025

Accepted: 20.11.2025

Date of publication: 05.04.2026