

# ЭТИЧНЫЙ ДИЗАЙН И ГУМАНИТАРНАЯ ЭКСПЕРТИЗА

*Антипов А. В.*

**Антипов Алексей Владимирович — к. филос. н., научный сотрудник Института философии Российской академии наук (ул. Гончарная, д. 12, стр. 1, Москва, 109240, Россия). Email: nelson02@yandex.ru.**

**Аннотация:** данная статья сосредоточена вокруг предположения о родственных положениях основных идей этического дизайна и гуманитарной экспертизы, которые сосредоточены на фундаменте этических норм. Этический дизайн определяется как возможность свободно совершать ценностный выбор при взаимодействии с технологическими новшествами, а также делается акцент на процессе совершения этого выбора. Выделяются положения, роднящие и связывающие этический дизайн и гуманитарную экспертизу. Первым положением является принципиальная незавершенность процесса диалога, в рамках которого происходит процесс формирования ценностного выбора и его реализация. Это позволяет не навязывать ценности в реализации готового продукта, а дает пользователю технологии выбор в процессе взаимодействия. Так происходит вовлечение конечного пользователя в реализацию технологических решений. Вторым положением обозначается выход понимания ответственности только за пределы безопасности, что позволяет говорить о более общих основаниях для принятия решения и вводить этическую нормативность в архитектуру

ру создаваемого новшества. В данном случае поднимаются вопросы о справедливости, предвзятости, прямом и косвенном ущербе, который может быть нанесен технологией. Третьим положением является акцент на прозрачности технологии, для чего необходимым элементом становится понимание процесса как проектирования, так и взаимодействия с пользователем. В рамках этичного дизайна актуализируется вопрос этических последствий, то есть задаются такие вопросы, которые касаются уровня доверия и подотчетности технологий, а также ее объяснимости. Как для гуманитарной экспертизы, так и для этичного дизайна важна возможность объяснения способа работы технологии для неспециалистов, поскольку это позволяет пользователям не взаимодействовать с технологией вслепую. Указанные вопросы рассматриваются в качестве того, что связывает этичный дизайн и гуманитарную экспертизу.

**Ключевые слова:** гуманитарная экспертиза, этичный дизайн, технологии, ответственность, справедливость, прозрачность, объяснимость

# ETHICAL DESIGN AND HUMANITARIAN EXPERTISE

*Aleksei V. Antipov*

**Aleksei V. Antipov — Cand. Sci. (Philosophy), researcher, Institute of Philosophy, Russian Academy of Sciences (Goncharnaya 12/1, Moscow, 109240, Russian Federation). Email: nelson02@yandex.ru.**

**Abstract.** This paper centers around the assumption that the core ideas of ethical design and humanitarian expertise are related and centered on a foundation of ethical standards. Ethical design is defined as the ability to freely make value choices when interacting with technological innovations, and the process of making these choices is emphasized. Provisions that relate and link ethical design and humanitarian expertise are highlighted. The first provision is the fundamental incompleteness of the dialog process, within which the process of forming a value choice and its realization takes place. This allows not to impose values in the realization of the finished product, but to give the user of technology a choice in the process of interaction. This is how the end user is involved in the realization of technological solutions. The second position indicates that the understanding of responsibility goes only beyond safety, which allows us to talk about more general grounds for decision-making and introduce ethical normativity into the architecture of the created innovation. In this case, questions are raised about fairness, bias, direct and indirect damage that can be caused by the technology. The third statement emphasizes a focus on the transparency of technolo-

gy, for which an understanding of both the design process and user interaction is a necessary element. Ethical design emphasizes the question of ethical implications, that is, it asks questions that address the level of trust and accountability of technology, as well as its explainability. For both humanitarian expertise and ethical design, the ability to explain how a technology works to non-specialists is important, as it ensures that users do not interact with the technology blindly. These issues are seen as what links ethical design and humanities expertise.

**Key words:** humanitarian expertise, ethical design, technology, responsibility, justice, transparency, explainability.

Современность характеризуется появлением большого количества технологических новшеств, изменяющих как саму жизнь, так и наши представления о ней. Происходит «все более плотное “обволакивание” человека наукой, его погружение в мир, проектируемый и обустраиваемый для него наукой и техникой»<sup>1</sup>. Возможно выделение двух крайних точек, между которыми разворачивается дискуссия: технологический пессимизм как стратегия выделения только негативных следствий и представление о том, что в привычном понимании структура жизни будет безвозвратно изменена и разрушена, а технологии виновны в обострении проблем современности, и технологический оптимизм, согласно которому технологии смогут решить многие возникающие затруднения. В рамках каждого из этих подходов является важным проведение всестороннего исследования, в том числе и гуманитарного. Только инженерный и позитивистский взгляд на технологии не способен описать и выделить возможные риски, к которым способны приводить технологии, поскольку указанные взгляды не учитывают специфику культурного и социального

---

<sup>1</sup> Юдин Б. Г. Технонаука, человек, общество: актуальность гуманитарной экспертизы // Век глобализации. № 2. 2008. С. 146-154.

ландшафта. Для выделения такого типа рисков разрабатывается гуманитарная экспертиза, которая способна также снизить градус алармизма технопессимистов, стремящихся видеть в технологиях только лишь негативные аспекты.

В рамках аналитического подхода предполагается рассмотреть общие положения, роднящие этический дизайн и гуманитарную экспертизу. Оба эти феномена рассматриваются как процессы, в рамках которых происходит выбор ценностей и их практическая реализация. В гуманитарной экспертизе выбор осуществляется через прогностическую деятельность и определение точек зрения, с которых рассматривается технологическое новшество. В этическом дизайне создается сама возможность осуществления выбора, что также связано с определением преобладающих ценностей в процессе разработки.

Контекст, который может быть охарактеризован как обращение к этике (или этический поворот), выражается в усилении внимания по отношению к этическим аспектам появляющихся технологий. На текущем этапе предполагается, что возможность этического выбора пользователя следует «вшить» непосредственно в технологию, сделать его составной частью дизайна и способ устройства. Подобное «вшивание» реализуется в двух моделях: модель привнесения, при которой возможность этического выбора привносится в изначально нейтральную технологию (например, предложение отказаться от бумажной квитанции за услуги ЖКХ в пользу электронной); и то, что можно назвать моделью реализации — выставление этического акцента на осуществляемом выборе (например, в приложениях для знакомств). Это определяется как этический дизайн, который, как можно предположить, реализуется на общих с гуманитарной экспертизой идеях, но находит свое отражение в конкретном устройстве или технологии.

Проведение гуманитарной экспертизы предполагает, что угрозы и риски, возникающие как следствие появления технологических новшеств, следует анализировать не только постфактум, но с помощью опережающего реагирования, «которое начинается

не после, а до того, как риск для человеческого потенциала становится очевидным»<sup>2</sup>. Гуманитарная экспертиза исходит из двух предпосылок: первая состоит в том, что любое новшество может считаться источником негативных последствий для человеческого потенциала; вторая выстраивается на предположении, что такого рода последствия оказываются непредвиденными вследствие недостаточного количества прилагаемых усилий. Ее целью является «определение, оценка и возможность корректирующих воздействий на факторы риска, которые актуально или потенциально несет в себе данная технология»<sup>3</sup>. Следует отметить, что не все технологии только вредоносны, но их следует разделять на три категории<sup>4</sup>:

1) технологии, относительно которых предостережения не оправдываются (технологии свободного доступа не уничтожили авторское право);

2) технологии, изменившие мир частично (появление цифровых носителей не полностью уничтожило сферу пленочных фильмов и фотокарточек);

3) технологии, глубоко проникшие и трансформировавшие окружающую реальность (такие как, например, Интернет).

Такое разделение позволяет рассматривать возникающие технологии не только как источник угроз, но и для того, чтобы убедиться в безопасности технологии и опережающим реагированием провести необходимые исследования посредством гипостазирования и прогнозирования.

Гипостазирование может быть рассмотрено как то, что различает этичный дизайн и гуманитарную экспертизу. Процедура применения этических нормативов к технологическим новшествам основывается на предположениях о потенциальном вреде,

---

<sup>2</sup> Ашмарин И. А., Юдин Б. Г. Человеческий потенциал: основы гуманитарной экспертизы // Человек: выход за пределы. М.: Прогресс-Традиция. 2018. С. 177.

<sup>3</sup> Там же. С. 188.

<sup>4</sup> Müller V. C. Ethics of Artificial Intelligence and Robotics // The Stanford Encyclopedia of Philosophy. Edward N. Zalta & Uri Nodelman (eds.). 2023. URL: <https://plato.stanford.edu/archives/fall2023/entries/ethics-ai/>

который технологии способны причинять. В то время как этичный дизайн исходит из того, что технологии уже не нейтральны, а содержат в себе определенное представление о способе функционирования. Однако данное замечание служит небольшим отступлением от основной темы исследования, которое состоит в выделении положений, роднящих этичный дизайн и гуманитарную экспертизу.

Этичный дизайн основан на возможности свободно устанавливать и формировать свой ценностный выбор в процессе взаимодействия с новшествами и технологиями. Предполагается выделение нескольких линий соприкосновения, в которых высвечиваются общие положения гуманитарной экспертизы и этичного дизайна. Первая линия состоит в принципиальной незавершенности процесса формирования этического выбора, который опирается на нормы. Для реализации этого «ценности не навязываются и не идентифицируются заранее, но остаются открытыми для настройки этических вариантов людьми»<sup>5</sup>. Это положение обосновывается тем, что невозможно обеспечить реализацию общемировых этических ценностей, но возможно создать процесс, который позволит приходить к решениям возникающих затруднений сознательно и процедурно строго. В гуманитарной экспертизе «основной целью во многих случаях является не конкретное решение, которое нередко чрезвычайно трудно найти, а процедура согласования интересов всех сторон»<sup>6</sup>. То же основание справедливо и для этичного дизайна: в качестве направления исследования этичный дизайн выстраивается на формировании процедур делиберации прежде всего между инженерами и представителями гуманитарной сферы, которые способны оценить

---

<sup>5</sup> Baldini G., Botterman M., Neisse R. et al. Ethical Design in the Internet of Things // *Sci Eng Ethics*. 24. 2018. P. 905–925

<sup>6</sup> Гребенищика Е. Г. Социогуманитарные контуры технонауки: актуальность гуманитарной экспертизы // *Социогуманитарные контуры технонауки: актуальность гуманитарной экспертизы* // *Знание. Понимание. Умение*. №. 1. 2018. С. 28–37.

пагубное влияние создаваемой технологии, а в широком смысле — согласование интересов разработчиков и пользователей.

Для реализации этих положений необходимо осуществить чувствительный к ценностям анализ, который выстраивается на выделении ценностей и их ранжировании в соответствии с господствующими представлениями. Основной акцент делается на процедурных особенностях, которые позволяют сделать процесс принятия решения прозрачным и воссоздаваемым. Следует признать, что не всегда разные группы исследователей приходят к единому результату относительно приоритетных ценностей и представлений о безопасности, но единый формат процедуры, состоящий в выполнении определенной последовательности шагов, способен сделать осуществляемый процесс соответствующим стандарту. Это предполагает и дает уверенность в том, что решение было принято профессионально, разные группы интересов были согласованы и соблюдены, а сам процесс прозрачен. Он «не гарантирует, что продукт не будет иметь негативных последствий, но он поможет снизить риски и предоставит профессионалам уверенность в том, что они действовали ответственно»<sup>7</sup>.

Все сказанное предполагает ответственный процесс проектирования. В него входит как рассмотрение последствий последовательно и профессиональными группами, так и включение в него процесса делиберации и согласования интересов. Это позволит если не избежать ценностных конфликтов, то снизить их негативные эффекты до минимальных. В ситуации моральных дилемм, когда оба варианта не удовлетворяют этическим требованиям, проверка решения (то есть соблюдение уровня прозрачности) и уверенность в том, что решения были выработаны надежными командами, позволяет минимизировать негативные эффекты.

Второй линией выступает приоритизация этических последствий и выход понимания ответственности за пределы только

---

<sup>7</sup> *Peters D., Vold K., Robinson D., Calvo R.* Responsible AI — Two Frameworks for Ethical Design Practice // *IEEE Transactions on Technology and Society*. Vol. 1. No. 1. 2020. P. 34-47.



безопасности. Вопросы, касающиеся справедливости, предвзятости, зависимости и косвенного ущерба, становятся предметом анализа в рамках разработки этического дизайна<sup>8</sup>. Связано это положение с исключением этического анализа из инженерной практики, что представляется неверным. Указывается, что, несмотря на то, что инженеры и разработчики и не должны выполнять работу социологов, философов и психологов, им необходимо сотрудничать и не пренебрегать специалистами гуманитарного профиля, «чтобы предвидеть и смягчать этические риски в качестве стандарта практики. Больше неприемлемо, чтобы технологии выходили в мир вслепую, оставляя других справляться с последствиями»<sup>9</sup>. Последнее положение очень знаменательно, поскольку указывает на принципиальный момент, связанный как с этическим дизайном, так и с гуманитарной экспертизой: необходимо реагировать на потенциальные опасности раньше, чем они проявили себя. Это не означает формирование алармистского дискурса, в рамках которого технологии демонизируются (в конце концов — технологии только используются людьми), а именно опережающее реагирование, способное на этапе проектирования предположить возможные проблемы.

Справедливость в случае этического дизайна понимается как равное распределение рисков, выгод и затрат без различения людей по индивидуальным характеристикам<sup>10</sup>, а также доступ к продуктам без приоритизации отдельных групп. Примером такого рода справедливого распределения может служить сетевой нейтралитет. Это принцип, согласно которому провайдеры не отдадут предпочтений сайтам разной направленности, то есть сайт букмекера должен получать такую скорость Интернета и такой же доступ, что и сайт по получению государственных услуг. В действительности такой принцип выполняется не всегда, и неко-

---

<sup>8</sup> Ibid

<sup>9</sup> Ibid.

<sup>10</sup> *Sanches P. et al.* HCI and affective health: Taking stock of a decade of studies and charting future research directions // Proc. CHI Conf. Human Factors Comput. Syst. 2019. P. 245.

торые государства ограничивают его действие. Этический дизайн предполагает использование ресурсов, особенно при их ограниченности, в равной мере, что может способствовать равномерному доступу к сервисам и продуктам, являющимся важными для разных категорий людей. Возможно поднять вопрос о различной ценности ресурсов (сайт получения государственных услуг более социально приемлем, чем сайт букмекера), однако ограничение сетевого нейтралитета создает пространство для потенциальных злоупотреблений.

Предвзятость характеризуется выделением одной группы среди всех остальных, для которой делается исключение из общего правила. Это могут быть индивидуальные характеристики (раса, пол, национальность и др.), на основании которых группа либо дискриминируется, либо ей отдается предпочтение. Предвзятый характер отдельных сервисов представляет собой угрозу для равномерного распределения делиберативных процедур и прав различных категорий людей, то есть способен приводить к политическим и социально негативным эффектам.

Проблема зависимости остро стоит в онлайн-сервисах, поскольку они призваны сделать так, чтобы пользователь проводил в них как можно больше времени. Так оказывается негативное влияние на пользователей, которые могут чувствовать себя хорошо в процессе использования, но их качество жизни от чрезмерной вовлеченности страдает. В контексте представлений об этическом дизайне предполагается соблюдение принципа благоприятного воздействия, даже если это способно наносить некоторый ущерб экономической рентабельности разработчика. В этом случае процедуры привлечения сторонних исследователей-гуманитариев способны добавить объективности в процесс оценки негативных воздействий, поскольку сами компании в процессе оценки собственного продукта не могут быть непредвзяты.

Косвенный ущерб связан с ущемлением определенных принципов (например, биоэтических, если говорить о продуктах, связанных с медициной), но напрямую на них не направлен. Рассмотрим на примере автономии. Зависимость от онлайн-игр и соци-

альных сетей способна приводить к снижению качества жизни, однако это также понижает степень автономии, так как уменьшает время, затрачиваемое на взаимодействие с другими, уход за собой и семьей и т. д., а также влияет на качество и свободу принимаемых решений, поскольку зависимость остро соотносит принимаемые решения со сформированной зависимостью нуждой. Если гуманитарная экспертиза призвана прогнозировать такого рода риски для автономии человека от использования определенных сервисов, то этический дизайн позволяет на уровне интерфейса, например, ограничивать время взаимодействия с программой, что связано с рисками для автономии. Другой проблемой является конфиденциальность пользовательских данных. Технологии шэринга, позволяющие обмениваться информацией, в том числе личной и зачастую чувствительной, также влияют на степень автономии, поскольку не позволяют пользователю полноценно контролировать способы и формы передачи и хранения его информации. Обсуждаемой в этом контексте проблемой является равенство возможностей, отсылающее к принципу справедливости. Разные люди исходя из уровня образования, интересов, опыта взаимодействия и др. на разном уровне взаимодействуют с появляющейся технологией.

Третьей линией является требование прозрачности. Для гуманитарной экспертизы прозрачность важна потому, что процесс выявления рисков и угроз становится возможным только при наличии определенного уровня прозрачности технологии, позволяющей взглянуть на технологическое новшество изнутри, а также в достаточной мере понять, каким образом технология встраивается в общий порядок вещей. В случае этического дизайна справедливо замечается, что прозрачность способна быть разной: прозрачность для инженера будет отличаться от прозрачности, предназначенной для пользователя<sup>11</sup>. Так, уровень прозрачности технологии влияет также на степень доверия к технологии в обществе. Крупные и резонансные события, такие как аварии

---

<sup>11</sup> Bryson J., Winfield A. Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems // Computer. Vol. 50. No. 5. 2017. P. 116-119.

с участием машин с автопилотом, вызывают волну недоверия и отрицания в обществе. Прозрачность связана с объяснимостью, более того, объяснимостью для обыкновенного пользователя, не обладающего специализированными знаниями. В гуманитарной экспертизе это требование обозначается как привлечение к экспертизе обычных людей, непрофессионалов (laypersons). Их участие предполагается для выстраивания конструктивного диалога с обществом, для которого объяснение должно быть открыто и понятно. Но объяснимость для конечного пользователя играет еще одну немаловажную роль: возможность получить ответ на вопрос, почему система поступает именно таким образом, а не иным. Взаимодействие с умной колонкой должно позволить человеку понять, почему выдаются именно такие, а не иные результаты. В случае нештатной ситуации объяснимость позволяет пользователю удостовериться, что решение интеллектуальной системой было принято на основании ясных и понятных непрофессионалу принципов.

Прозрачность влияет и на подотчетность. Правильность работы системы обеспечивает возможность контролировать ее. В условиях несовершенства интеллектуальных систем (а даже при их совершенности контроль и последнее слово за человеком, как можно предположить, усилят доверие) возможность останавливать работу или выбирать нужное действие позволяют также увеличивать безопасность работы системы.

Как указывают исследователи этичного дизайна, он должен быть основан на этической конструкции, которая имеет следующие функции: (а) возможность обеспечивать контроль над сбором и распространением данных или услуг, связанных с пользователем. Другими словами, они направлены на поддержку этических возможностей людей, таких как субъектность, осведомленность и рефлексивность (требующая прозрачности в том, как собираются и распространяются данные); (б) способность обеспечивать соблюдение различных культурных правил с течением времени (например, культурные или нормативные изменения) или пространства (например, разные страны); (с) способность

поддерживать динамические контексты (например, дом, офис) и (г) способность воспринимать, идентифицировать и поддерживать отношения, которые требуют этического выбора<sup>12</sup>. Эти положения выводят этический дизайн за пределы только гуманитарной экспертизы и позволяют увидеть в нем способ воплощения идей о защите безопасности, приватности, уважении автономии, которые могут быть реализованы в рамках конкретных интеллектуальных систем.

Таким образом, понимание этического дизайна как формирования процесса, в рамках которого происходит формирование ценностного выбора, а также делается упор на континуальности, служит доказательством родственных отношений между этическим дизайном и основными идеями гуманитарной экспертизы.

## Литература

*Ашмарин И.А., Юдин Б. Г.* Человеческий потенциал: основы гуманитарной экспертизы // Человек: выход за пределы. М.: Прогресс-Традиция. 2018. С. 176-193.

*Гребенщикова Е. Г.* Социогуманитарные контуры технонауки: актуальность гуманитарной экспертизы // Социогуманитарные контуры технонауки: актуальность гуманитарной экспертизы // Знание. Понимание. Умение. №. 1. 2018. С. 28-37.

*Юдин Б. Г.* Технонаука, человек, общество: актуальность гуманитарной экспертизы // Век глобализации. № 2. 2008. С. 146-154.

*Baldini G., Botterman M., Neisse R. et al.* Ethical Design in the Internet of Things // *Sci Eng Ethics*. 24. 2018. P. 905–925

*Bryson J., Winfield A.* Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems // *Computer*. Vol. 50. No. 5. 2017. P. 116-119.

*Müller V. C.* Ethics of Artificial Intelligence and Robotics // *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta & Uri Nodelman (eds.). 2023. URL: <https://plato.stanford.edu/archives/fall2023/entries/ethics-ai/>

---

<sup>12</sup> *Baldini G., Botterman M., Neisse R. et al.* Ethical Design in the Internet of Things // *Sci Eng Ethics*. 24. 2018. P. 905–925.

*Peters D., Vold K., Robinson D., Calvo R.* Responsible AI — Two Frameworks for Ethical Design Practice // IEEE Transactions on Technology and Society. Vol. 1. No. 1. 2020. P. 34-47.

*Sanchez P. et al.* HCI and affective health: Taking stock of a decade of studies and charting future research directions // Proc. CHIConf. HumanFactorsComput. Syst. 2019. P. 245.