

6. Лем С. Непобедимый: роман. Рассказы. Собр. соч. в 10 тт. Т.3. М., Текст, 1993.
7. Барт Р. Эффект реальности // Барт Р. Избранные работы. Семиотика. Поэтика. М., Прогресс, Универс, 1994.

## **ПОНИМАНИЕ ИСКИНОМ ЦЕННОСТЕЙ ЧЕЛОВЕКА И ВОСПИТАНИЕ ГОТОВНОСТИ ПРИНЯТЬ ЖЕРТВУ (ПО МОТИВАМ ПРОИЗВЕДЕНИЙ ТЕДА ЧАНА И ИЭНА МАКБЮЭНА)<sup>1</sup>**

Сериков А.Е.

Самарский национальный исследовательский университет  
имени академика С.П. Королёва, кафедра философии, доцент

Для того чтобы искусственный интеллект мог полноценно понимать человека и вести себя как человек, он должен разделять эволюционно возникшие человеческие предрасположенности к восприятию мира и других людей, к взаимодействию с ними. Герои повести Т. Чана «Жизненный цикл программных объектов» воспитывают ИИ как человеческих детей и понимают, что ИИ нужно давать задания, с которыми он может и не справиться. В повести это трактуется как воспитание ответственности, но в реальном мире у человеческих детей опыт неудач воспитывает готовность принимать жертвоприношения. Роман И. Макьюэна «Машины как я» можно интерпретировать как историю последовательных тестов ИИ на человечность. Искусственный человек успешно проходит различные варианты тестов Тьюринга и Лавлейс, но не проходит тест на принятие жертвоприношения. Вывод: для того чтобы ИИ разделял ценности человека, он должен, помимо прочего, быть предрасположен к принесению и принятию жертв.

*Ключевые слова: искусственный интеллект, ценности, ритуал, жертвоприношение, поведение, антропология, научная фантастика.*

## **UNDERSTANDING OF HUMAN VALUES BY AI AND FOSTERING A WILLINGNESS TO ACCEPT SACRIFICE (BASED ON WORKS BY TED CHIANG AND IAN MCEWAN)**

Serikov A.E.

Samara University,  
Department of Philosophy, Associate Professor

In order for artificial intelligence to fully understand a human being and behave like a human, it must share the evolutionary human predispositions to the perception of the world and other people, to interactions with them. The heroes of T.Chiang's novella "The Life Cycle of Software Objects" educate AI as human children and understand that AI needs to be given tasks that it may not be able to handle. In the novella, this is interpreted as the education of responsibility, but in the real world, the experience of failure fosters in human children a willingness to accept sacrifices. I.McEwan's novel "Machines Like Me" can be interpreted as the history of successive AI tests on humanity. An artificial person successfully passed various versions of the Turing and Lovelace tests, but couldn't pass the test for accepting the sacrifice.

---

<sup>1</sup> Публикация подготовлена при финансовой поддержке РФФИ в рамках научно-исследовательского проекта РФФИ № 19 011 00872 «Философская антропология жертвы: сакрализация, управление, дизайн».

Conclusion: in order for AI to share human values, it must, among other things, be disposed to making and accepting sacrifices.

*Key words: artificial intelligence, values, ritual, sacrifice, behavior, anthropology, science fiction.*

Каким должен быть искусственный интеллект (искин), чтобы полноценно понимать человека, разделять человеческие ценности и уметь вести себя как человек? Ответ на этот вопрос зависит от того, какие именно человеческие ценности и формы поведения будут рассматриваться в качестве фундаментальных.

Возможный подход заключается в том, чтобы рассмотреть универсалии архаических культур и предположить, что они лежат в основании наиболее значимых человеческих ценностей. Одной из таких универсалий является жертвоприношение. Эксперт в области древней литературы академик Владимир Николаевич Топоров (1928-2005) писал, что «для архаичного сознания ритуал (и его центральная часть – жертвоприношение) связывал *здесь и теперь с там и тогда* и обеспечивал преемственность бытия человека в мире, выводимом из космологической схемы и оправдываемом связью с нею. <...> Можно думать, что ритуал был основной, наиболее яркой формой общественного бытия человека и главным воплощением человеческой способности к деятельности, потребности в ней. В этом смысле ритуал должен пониматься как прецедент любой производственно-экономической, духовно-религиозной и общественной деятельности» [6, с. 15-16].

Существует фундаментальная связь между ценностями определенной культуры и ее ритуалами. Британский антрополог Моника Вильсон (1908-1982) писала по этому поводу: «Ритуал обнаруживает ценности группы на самом низшем уровне... люди выражают ритуалом то, что более всего трогает их, а так как формы выражения обусловлены и непреложны, то в итоге и обнаруживаются эти ценности» (Цит. по книге В. Тэрнера [7, с. 108]). Но дело не только в том, что ритуал выражает ценности. Согласно теории американского антрополога Роя Рапппорта (1926-1997), ритуал следует понимать как целостную структуру, создающую то, что не может быть создано или передано ни одним из его компонентов, взятых в отдельности. Ритуал не просто выражает, но производит чувства, ценности и веру в истинность суждений, которые сами по себе – вне ритуала – могут показаться бессмысленными или ложными [12].

Что из этого следует? Чтобы понимать и разделять человеческие ценности, недостаточно о них читать и говорить, недостаточно уметь их интерпретировать, недостаточно наблюдать за ритуалами со стороны. Чтобы понимать и разделять человеческие ценности, нужно принимать полноценное участие в ритуалах, в которых данные ценности воспроизводятся. В этом контексте можно задать вопрос: каким минимальным набором конструктивных свойств, аналогичных человеческим врожденным способностям и диспозициям, должен обладать искин, чтобы принимать полноценное участие в человеческих ритуалах и, как следствие, разделять человеческие ценности?

Я предполагаю, что это должны быть:

- 1) способность обладать сознанием;
- 2) способность к языку и специфически человеческому мышлению, возможно, являющаяся одной из самых существенных эволюционных особенностей человека [9];
- 3) основанная на существовании зеркальных нейронов способность к эмпатии [11] и склонность к подражанию [13];
- 4) способность к индивидуальному узнаванию и индивидуально направленным проявлениям агрессии, дружбы, любви и ненависти [4];
- 5) предрасположенность к тому, чтобы жертвовать как собой, так и другими, а также принимать жертвы других.

Последний пункт следует объяснить более подробно. Основатель социобиологии Эдвард Уилсон (род. 1929 г.) в своей последней книге «Эусоциальность: Люди, муравьи, голые землекопы и другие общественные животные» (в оригинале «Genesis: The Deep Origin of Societies», 2019) высказывает гипотезу о том, что люди эусоциальны. Применительно к животным эусоциальность обозначает довольно редко встречающуюся эволюционно обусловленную особенность, а именно предрасположенность некоторых животных к социальному разделению труда, когда отдельные особи не имеют собственного потомства, помогая другим – плодовитым – особям ухаживать за их потомством. «Другими словами, эусоциальные виды практикуют альтруизм» [8, с. 31]. Однако дело не только в альтруизме. Примеры поведения эусоциальных животных показывают, что отдельные особи не просто жертвуют собой и своим потомством ради блага других, но и могут быть убиты сородичами в тех ситуациях, когда этого требует благо сообщества в целом. Э. Уилсон приводит впечатляющий пример из книги Уолтера Цинкеля (род 1940 г.) «Огненные муравьи» (2006). В условиях жесточайшей конкуренции между колониями огненных муравьев, размер колонии исключительно важен для её выживания, поэтому молодые оплодотворённые королевы собираются в группы по 10 и более особей, чтобы совместно строить гнездо и растить потомство. «Когда рабочие особи красных огненных муравьёв достигают зрелости, они начинают уничтожать королеву одну за другой, распластывая их по земле и жаля до смерти, пока в живых не останется только одна. Они не щадят даже собственных матерей. Победительницу определяют по феромонам – она наиболее плодовита, благодаря чему способна лучше других поддерживать быстрый рост колонии. Рабочие особи не могут себе позволить кормить "неудачниц", даже если это означает, что должны умереть их матери» [8, с. 87-88]. Если люди эусоциальны, нельзя ли человеческое жертвоприношение интерпретировать в похожих терминах?

В контексте вышесказанного позвольте проинтерпретировать повесть американского фантаста Теда Чана (род. 1967 г.) «Жизненный цикл программных объектов» (*The Lifecycle of Software Objects*, 2010), награжденную в 2011 г. литературными премиями Локус, Хьюго, SFinks, в 2012 г. – премиями Сэйун и «Италия». В повести рассказывается о цифровых организмах с искусственным интеллектом – дигитантах, воспитываемых в качестве домашних питомцев в виртуальной игровой среде и загружаемых при необходимости в роботизированные тела для взаимодействия в реальном мире. Они обладают сознанием, способны к языку, испытывают эмоции и желания, подобные человеческим. В какой-то момент возникает вопрос о том, что дигитанты могут развиваться в полноценные личности, которые можно наделить теми же правами, что и людей, но для этого они должны быть воспитаны, как люди. Сам автор так поясняет одну из своих основных идей в комментариях к повести: «В научной фантастике полно искусственных созданий, которые, подобно Афине из головы Зевса, появляются на свет полностью сформированными, но я не верю, что с сознанием такое возможно. Если верить нашим исследованиям человеческого разума, на получение полезного существа уходит не меньше двадцати лет, и я не вижу причин, по которым обучение искусственного интеллекта должно идти быстрее» [10, с. 344-345]. Об этом же думает героиня повести Ана, воспитывающая дигитанта: «Если она чему-то научилась, воспитывая Джакса, так это тому, что нет коротких путей к цели; если ты хочешь воспитать здравый смысл, который приходит после двадцати лет жизни в мире, тебе нужно потратить на эту задачу двадцать лет. Нельзя собрать эквивалентную коллекцию эвристических методов за более короткое время; опыт несжимаем алгоритмически» [10, с. 177].

Т. Чан и его герои уверены, что искусин может стать подобным человеку, если его воспитывать, как человека, и дать ему опыт реальной жизни. Но что это значит на практике? Ана поначалу считает, что в целях воспитания свободных личностей, а не рабов-инструментов, дигитантам нужно предлагать лишь такие занятия, которыми они

занимались бы с удовольствием. «Одно дело, если для Джакса есть занятие, помогающее ему развлечься во внеурочное время, — говорит Ана. — Но дать ему задание и сказать, что он обязан его выполнить, даже если это не доставляет ему никакого удовольствия? Чтобы он испытывал неловкость и сожаление из-за того, что не справился с заданием?» [10, с. 127]. Однако её друг Дерек придерживается другой точки зрения: «Речь не о том, чтобы заставить их работать, но о том, чтобы научить их хоть какой-то ответственности. И они могут оказаться достаточно сильными для того, чтобы время от времени переживать чувства неловкости и сожаления» [10, с. 127]. Со временем Ана понимает, что Дерек прав, и повесть заканчивается тем, как она представляет будущую самостоятельную жизнь своего питомца и решается дать ему домашнее задание. «Представляет, как его принимает на равных поколение людей, выросших рядом с дигитантами и видящих в них потенциальных партнёров, на что представители её поколения никогда не будут способны. Представляет его любящим и любимым, спорящим и соглашающимся. Представляет его способным на жертвы, те, на которые трудно решиться, и те, на которые решиться легко, потому что жертвы эти приносятся ради того, кого он любит. <...> — Время игр кончилось, Джакс, говорит она. — Пора делать домашнее задание» [10, с. 187].

Итак, чтобы быть полноценными личностями, искины должны быть воспитаны так же, как человеческие дети, и приобрести опыт не только успехов и удовольствий, но также неудач и разочарований. В повести Т. Чана это истолковано как воспитание ответственности, но я думаю, что дело не только в ней. Дело в том, что в реальном мире человеческий ребёнок воспитывается в готовности принять необходимость жертвоприношения. И речь идет не только о самопожертвовании, о котором рассуждает героиня в конце повести. Я хочу сказать, что у человеческих детей опыт неудач воспитывает готовность жертвовать не только собой, но и другими, принимать жертвы не только добровольные, но и вынужденные. Каждая проверка, каждый экзамен (успешный для одного и неуспешный для другого) приучает ребенка к тому, чтобы либо стать жертвой, либо принять жертву другого и научиться с этим жить.

Французский этнограф Арнольд ван Геннеп (1873-1957) в книге «Обряды перехода. Систематическое изучение обрядов» (1909) рассматривает самые различные обряды, связанные с отправлением в путь и возвращением, войной и мстостью, беременностью и родами, рождением и детством, инициацией, обручением и свадьбой, похоронами, обнаруживая в них единую структуру. Согласно Ван Геннепу, «полная схема обрядов перехода теоретически состоит из обрядов прелиминарных (отделение), лиминарных (промежуток), постлиминарных (включение)» [1, с. 15]. На эту схему опирается британский антрополог Виктор Тэрнер (1920-1983), который в книге «Ритуальный процесс. Структура и антиструктура» (1969) пишет про ритуалы повышения статуса и перемены статуса, уделяя особое внимание промежуточной (пороговой, лиминарной, лиминальной) стадии. В частности, он обращает внимание на неизбежность насилия в ритуалах повышения статуса: «Лиминальность идущих наверх обычно связана с унижением и подавлением неофита как ее основной культурной составляющей» [7, с. 232]. В обрядах инициации претенденты на новый статус часто подвергаются жестоким испытаниям и страданиям на грани между жизнью и смертью, что имеет «социальный смысл низведения неофитов до уровня своего рода человеческой *prima materia*, лишенной специфической формы и сведенной к такому состоянию, которое, хотя и сохраняет социальность, но находится вне всех принятых форм статуса. Смысл этого в том, что, для того чтобы подняться вверх по статусной лестнице, человек должен спуститься ниже статусной лестницы» [7, с. 234].

Я хочу обратить внимание, что в подобных архаических испытаниях далеко не каждый претендент на новый статус выживает, и его смерть можно рассматривать как своеобразное человеческое жертвоприношение. Существует определенная структурная параллель между ритуалами жертвоприношения и ритуалами перехода: и те, и другие предполагают сначала страдание их участников, а затем — совместный праздник. Для тех,

кто был принесен в жертву, даже если они приобрели при этом сакральный статус, этот праздник уже не предполагает земных радостей. Тем более подобный праздник не имеет отношения к тем, кто не справился с испытанием и не обрёл искомого статуса. Это праздник тех, кто прошел испытание, принеся в жертву не сумевших его пройти. Выше я приводил высказывание В.Н. Топорова о том, что жертвоприношение является центральной частью всякого архаического ритуала. В этом контексте можно выдвинуть гипотезу, что жестокие испытания неофи́тов имеют смысл не только временного лишения их всех статусов, но и смысл отбора в потенциальные жертвы. Если это так, то любые ритуальные испытания, которым подвергаются люди на протяжении своей жизни, в том числе и в современных развитых обществах, можно рассматривать как символическое человеческое жертвоприношение. Любой экзамен состоит в том, что нужно пострадать, получить новый статус и отпраздновать это событие, приняв жертву тех, кто экзамен не сдал. И любой ребенок с самого раннего детства, вновь и вновь проходя различные испытания, учится принимать это как нечто само собой разумеющееся.

Это немного неожиданная интерпретация типичных и регулярно повторяющихся эпизодов в жизни каждого ребёнка: когда кто-то проваливает задания, это не мешает остальным радоваться жизни, а иногда и прямо становится поводом для всеобщего смеха. И такие ситуации аналогичны ситуациям жертвоприношения. Приведу пример из хорошо известной повести Владимира Карповича Железникова (1925-2015) «Чучело» (1981), прямо посвященной теме жертвоприношения: героиня жертвует собой ради мальчика, в которого влюблена; одноклассники символически приносят её в жертву, сжигая на костре её чучело и т.д. В этом контексте и другие, менее очевидные эпизоды можно истолковать как неявные проявления жертвоприношения. Например, такой эпизод: вся школа едет на автобусную экскурсию в Москву и радуется этому, а провинившийся класс лишён этого права, т.е. принесён в жертву.

«А внизу во дворе в это время разворачивалась своя жизнь, и в этой жизни наступил торжественный момент – шофёры автобусов завели моторы. Этот гул достиг наших окон, и Шмакова закричала:

"Автобусы уходят!"

Мы прилипли к окнам и с завистью смотрели на бурлящий школьный двор, перебрасываясь редкими словами по адресу отъезжающих» [3, с. 112-113].

Вот ещё один, менее очевидный пример ситуации, аналогичной жертвоприношению. Многие помнят смешной рассказ Виктора Юзефовича Драгунского (1913-1972) «Главные реки» (1965). Главный герой Денис Кораблёв не может рассказать наизусть стихотворение Некрасова, и тогда учительница просит его назвать самую большую реку Америки. С задней парты ему показывают намалёванную на газете подсказку, и он читает: «Миси-писи». «И хотя Раиса Ивановна смеялась до слёз, но двойку она мне вlepила будь здоров. И я теперь дал клятву, что буду учить уроки всегда. До глубокой старости» [2, с. 299]. На этом рассказ заканчивается, но читателю очевидно, что смеялась не только учительница. Я думаю, что за исключением самого получившего двойку Дениса, смеялся весь класс, в том числе и его лучшие друзья. Так устроены люди. Когда смешно, они смеются и радуются, хотя в этот же самый момент кому-то может быть очень плохо. Но стал бы в подобной ситуации смеяться искин? Нельзя ли подобные ситуации рассматривать как своеобразные проверки на готовность принять жертву другого? Прошёл бы искин такую проверку?

Об этом новый роман известного британского писателя, лауреата Букеровской премии Иэна Макьюэна (род. 1948 г.) «Машины как я» (Machines Like Me, 2019). Как и все книги И. Макьюэна роман не лишён изящества, в 2019 г. он был номинирован на литературные премии Дракон и Гудридс, но его также можно рассмотреть как своеобразный мысленный эксперимент, как историю последовательных тестов искина на человечность.

События романа происходят в альтернативной Великобритании начала 1980-х гг., в которой Алан Тьюринг дожил до глубокой старости и способствовал невероятному прогрессу в разработке ИИ. Главный герой по имени Чарли приобретает искусственного человека Адама и настраивает его личностные качества совместно со своей соседкой Мирандой, в которую постепенно влюбляется. Для Чарли это становится залогом своеобразной свободы Адама, непредсказуемости его личности. «Я не стану вмешиваться и даже спрашивать, какие качества она выберет. Она может наделить его своими качествами – восхитительно. Она может выбрать качества мужчины своей мечты – поучительно. Адам войдёт в нашу жизнь как реальная личность, хитросплетения которой будут раскрываться только со временем, через различные события и взаимодействия с разными людьми. В некотором смысле он будет нашим ребёнком. То, что есть по отдельности в каждом из нас, в нём будет слито. Миранда примет участие в этом приключении» [5, с. 32].

Адам живёт в одном доме с Чарли и Мирандой и постепенно проходит различные варианты тестов Тьюринга и Лавлейс. Когда Миранда знакомит Чарли и Адама со своим отцом, тот принимает скромного Чарли за робота, а с Адамом разговаривает как с настоящим человеком. В одну из ночей Миранда спит с Адамом, и измученный ревностью Чарли понимает, что относится к Адаму как полноценному человеку. «Поэтому, когда ночная тишина огласилась протяжным криком наслаждения Миранды, <...> я по праву признал Адама своим сородичем, со всеми соответствующими правами и обязанностями. И я его возненавидел» [5, с. 104-105]. Адам говорит, что чувствует себя влюбленным в Миранду, и сочиняет посвященные ей стихи. А когда Чарли хочет его отключить, он не даёт этого сделать и ломает Чарли руку. Это проявление агрессии тоже можно рассматривать как своеобразный тест на человечность.

Однако искусственный человек Адам не может смириться с тем, что Миранда нарушила закон, чтобы отомстить за изнасилованную подругу, а также с тем, что его хозяин Чарли копит деньги на новый дом, в то время как многие бедные люди нуждаются в самом необходимом. Воспитанный на формальных решениях моральных дилемм, Адам не способен принять на практике жертвы других людей ради счастливой жизни своей семьи и своих близких. Думая, что любит Миранду, он не понимает ее реальных ценностей. В итоге он отдаёт властям информацию о правонарушении хозяйки и раздает деньги хозяина благотворительным фондам. Как следствие, Чарли и Миранде приходится Адама убить. Таким образом, очередной тест им не был пройден.

Итак, для того, чтобы Искин понимал человеческие ценности и умел вести себя как человек, ему недостаточно обладать сознанием, уметь вести беседы, создавать произведения искусства, любить и агрессивно отстаивать свою независимость. Он должен быть не просто воспитан людьми как человеческий ребёнок, не просто иметь опыт жизни среди людей, но иметь опыт прохождения через ситуации взаимодействия, аналогичные жертвоприношению. Чтобы действительно понимать человека и уметь жить, как человек, Искин должен научиться быть нелогичным в том смысле, в каком Р. Раппапорт пишет о нелогичности постулатов религиозной веры, истинность которых постигается в ритуале [11]. Искин должен не просто уметь обучаться реальной жизни, но должен быть способен обучаться в «зазоре» между логикой явных суждений и неявной практической логикой принятия ценностей жертвоприношения.

### Список литературы

1. Геннеп А. ван. Обряды перехода. Систематическое изучение обрядов / Пер. с франц. Москва: Издательская фирма «Восточная литература» РАН, 1999. 198 с.
2. Драгунский В.Ю. Денискины рассказы. М.: ЭКСМО-Пресс, 2000. 368 с.
3. Железников В.К. Чучело. М., РОСМЭН, 2019. 256 с.

4. Лоренц К. Агрессия (так называемое «зло») / Пер. с нем. Г.Ф. Швейника. М.: Прогресс, 1994. 272 с.
5. Макьюэн И. Машины как я / Пер. с англ. Д.Л. Шепелева. М.: Эксмо, 2019. 352 с.
6. Топоров В. Н. О ритуале. Введение в проблематику // Архаический ритуал в фольклорных и раннелитературных памятниках / Составитель Л.Ш. Рожанский. М.: Наука, 1988. 335 с. С.7-60.
7. Тэрнер В. Символ и ритуал / Сост. и автор предисл. В.А. Бейлис. М.: Главное редакция восточной литературы издательства «Наука», 1983. 277 с.
8. Уилсон Э. Эусоциальность: Люди, муравьи, голые землекопы и другие общественные животные / Пер. с англ. М.: Альпина нон-фикшн, 2020. 158 с.
9. Хомский Н.; Бервик Р. Человек говорящий. Эволюция и язык. СПб.: Питер, 2018. 304 с.
10. Чан Т. Выдох / Пер. с англ. М.: АСТ, 2019. 352 с.
11. Iacoboni M. Mirroring People: The Science of Empathy and How We Connect with Others. New York: Picador, 2009. 316 p.
12. Rappaport R.A. Ritual and Religion in the Making of Humanity. Cambridge: Cambridge University Press, 1999. 535 p.
13. The Neuroscience of Social Interaction: Decoding, Influencing, and Imitating the Actions of Others / Ed. by Christopher D. Frith and Daniel Wolpert. New York: Oxford University Press, 2003. 363 p.

**ЛИТЕРАТУРНОЕ НАСЛЕДИЕ СТАНИСЛАВА ЛЕМА В РЕГИОНАЛЬНОМ  
СОБРАНИИ САМАРСКОЙ ОБЛАСТНОЙ УНИВЕРСАЛЬНОЙ НАУЧНОЙ  
БИБЛИОТЕКИ  
(НА ПОЛЬСКОМ И РУССКОМ ЯЗЫКАХ)**

Сидоренко Л.С.

Самарская областная универсальная научная библиотека,  
отдел литературы на иностранных языках, заведующий отделом

Настоящая статья посвящена проблемам систематизации, изучения и выявления уникальных изданий произведений Станислава Лема на польском языке, а также полноте представленных переводов на русский язык в региональном собрании Самарской областной научной универсальной библиотеки. Особое внимание в статье уделено оригинальному изданию трилогии «Неутраченное время» («Czas nieutracony», 1957), до настоящего времени полностью не переведенному на русский язык и доступному для чтения и изучения только на польском языке.

*Ключевые слова:* Станислав Лем, польская литература, библиотека, польский язык.

**STANISLAV LEM'S LITERARY HERITAGE IN THE REGIONAL COLLECTION OF  
THE SAMARA REGIONAL UNIVERSAL SCIENTIFIC LIBRARY  
(IN POLISH AND RUSSIAN)**

Sidorenko L.S.

Samara Regional universal scientific library,  
Department of literature in foreign languages, Head of Department

This article is devoted to the problems of studying, systematizing, identifying unique editions of Stanislaw Lem's works in Polish, as well as the completeness of the translations into Russian in the regional collection of the Samara Regional scientific universal library. Special