НОВАЯ ЭРА НАУКИ И ЗНАНИЙ
МЕЖДУНАРОДНЫЙ НАУЧНЫЙ ЖУРНАЛ
global-science-review.com
editor@global-science-review.com

ЭКОНОМИКА ДОВЕРИЯ В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: КАК ЦИФРОВЫЕ ТЕХНОЛОГИИ МЕНЯЮТ ПОВЕДЕНИЕ ПОТРЕБИТЕЛЕЙ

Адзова Мадинаа Абдулазиз кизи

“студент 4 курса”

Факультет Экономики,

Кафедра Экономика транспорта, http://orcid.org/ORCID:0009-0007-9785-1117

Ташкентский Государственный Транспортный Университет , Ташкент,

Узбекистан.

*e-mail:**madinaadizova59@gmail.com***

Аннотация. Статья посвящена трансформации доверия как экономического ресурса в условиях массового внедрения систем искусственного интеллекта (ИИ) в потребительские рынки. Цель исследования — выявить механизмы, определяющие доверие и недоверие потребителей к алгоритмическим решениям при совершении покупок, в потребительском кредитовании и при выборе услуг. Методологической основой работы является интегративный обзор рецензируемых публикаций в области поведенческой экономики, маркетинга, финансов, информационных систем и инженерной психологии, дополненный анализом регуляторных документов. Показано, что доверие к ИИ носит двухуровневый характер: ожидания доброжелательности и честности, неприменимые к алгоритму, переносятся на оператора системы и институциональную среду. Установлено, что противоречие между эффектами алгоритмического отторжения и алгоритмического признания снимается при учёте характера задачи, опыта взаимодействия с системой и возможности контроля над её решениями. Сравнительный анализ трёх сфер демонстрирует, что по мере роста ставок решения механизм формирования доверия смещается от приобретённого

опыта к институциональным гарантиям, а ключевыми факторами становятся объяснимость, справедливость и право на оспаривание. Предложена концептуальная модель калиброванного доверия потребителей к ИИ и сформулированы семь исследовательских предложений для последующей эмпирической проверки. Результаты имеют значение для компаний, применяющих ИИ во взаимодействии с клиентами, и для регуляторов цифровых рынков.

Ключевые слова: экономика доверия, искусственный интеллект, поведение потребителей, алгоритмическое отторжение, калиброванное доверие, объяснимый ИИ, кредитный скоринг, рекомендательные системы.

Abstract. This article examines how trust, as an economic resource, is being transformed by the widespread deployment of artificial intelligence (AI) in consumer markets. The study aims to identify the mechanisms that shape consumer trust and distrust in algorithmic decisions in retail purchasing, consumer lending, and service selection. The methodology is an integrative review of peer-reviewed research in behavioural economics, marketing, finance, information systems, and human factors, complemented by an analysis of regulatory documents. The findings show that trust in AI is two-tiered: expectations of benevolence and integrity, which cannot apply to an algorithm itself, are transferred to the system's operator and to the institutional environment. The apparent contradiction between algorithm aversion and algorithm appreciation is resolved once task type, experience with the system, and the possibility of controlling its outputs are taken into account. A comparison of the three domains demonstrates that as decision stakes rise, the basis of trust shifts from accumulated experience to institutional safeguards, while explainability, fairness, and the right to contest decisions become decisive. The article proposes a conceptual model of calibrated consumer trust in AI and formulates seven research propositions for future empirical testing. The results are relevant to firms using AI in customer interactions and to regulators of digital markets.

Keywords: economics of trust, artificial intelligence, consumer behaviour, algorithm aversion, calibrated trust, explainable AI, credit scoring, recommender systems.

1. Введение

Доверие принадлежит к числу тех экономических категорий, значение которых становится очевидным главным образом в момент их утраты. К. Эрроу (Arrow, 1972) отмечал, что практически каждая коммерческая сделка содержит в себе элемент доверия, а многие проявления экономической отсталости могут быть объяснены его недостатком. Последующие исследования подтвердили, что уровень межличностного и институционального доверия связан с темпами экономического роста (Zak & Knack, 2001), а его дефицит порождает проблему асимметрии информации, способную разрушить рынок, как это показано в классической модели «рынка лимонов» Дж. Акерлофа (Akerlof, 1970). Доверие снижает транзакционные издержки, позволяя сторонам экономить ресурсы на контроле, проверке и страховании рисков оппортунистического поведения (Williamson, 1993).

Цель настоящей статьи состоит в интеграции разрозненных результатов в единую концептуальную рамку, объясняющую, почему потребители доверяют или не доверяют решениям ИИ в трёх ключевых сферах: при совершении покупок, в потребительском кредитовании и при выборе услуг. В соответствии с этой целью сформулированы следующие исследовательские вопросы.

RQ1. Какие механизмы лежат в основе доверия и недоверия потребителей к алгоритмическим решениям?

RQ2. Как различаются детерминанты доверия к ИИ в сферах розничных покупок, кредитования и услуг?

RQ3. Какие условия обеспечивают формирование калиброванного, то есть соответствующего реальной надёжности системы, доверия к ИИ?

Вклад работы заключается в трёх аспектах. Во-первых, предложена трактовка доверия к ИИ как двухуровневого феномена, в котором доверие к технологии опосредовано доверием к её оператору и институциональной среде. Во-вторых, проведён сравнительный анализ трёх потребительских сфер, различающихся величиной ставок и характером решений. В-третьих, разработана концептуальная модель калиброванного доверия и сформулированы исследовательские предложения, пригодные для последующей эмпирической проверки.

Статья организована следующим образом. Во втором разделе рассматриваются теоретические основания исследования. Третий раздел описывает методологию. В четвёртом разделе представлены результаты синтеза литературы по трём сферам. Пятый раздел содержит концептуальную модель и исследовательские предложения. В шестом разделе обсуждаются теоретические и практические импликации, в седьмом — ограничения и направления будущих исследований. Заключение подводит итоги работы.

2. Теоретические основания

В экономической теории доверие рассматривается с нескольких позиций. В институциональной традиции оно трактуется как механизм снижения неопределённости и транзакционных издержек. О. Уильямсон (Williamson, 1993) предложил разграничивать «калькулятивное» доверие, основанное на расчёте выгод и рисков, и доверие в собственном смысле слова, не сводимое к рациональному расчёту. Социологическая традиция, восходящая к Н. Луману (Luhmann, 1979), рассматривает доверие как способ редукции социальной сложности: доверяя, индивид принимает решение, не располагая полной информацией, и тем самым делает возможным действие в условиях неопределённости.

В междисциплинарном обзоре Д. Руссо и соавторов (Rousseau et al., 1998) доверие определено как психологическое состояние, включающее

намерение принять уязвимость на основе позитивных ожиданий относительно намерений или поведения другой стороны. Это определение подчёркивает два ключевых элемента, значимых для анализа ИИ: уязвимость доверителя и ожидания относительно поведения доверенной стороны.

3. Методология

Исследование выполнено в формате интегративного обзора литературы, который, в отличие от систематического обзора с метаанализом, допускает синтез работ, различающихся по методам, дисциплинарной принадлежности и уровню анализа, и ориентирован на построение концептуальных рамок. Выбор данного формата обусловлен междисциплинарным характером проблемы и неоднородностью эмпирических данных.

Поиск источников осуществлялся в базах Scopus, Web of Science и Google Scholar с использованием комбинаций ключевых слов «trust in AI», «algorithm aversion», «algorithm appreciation», «consumer trust», «algorithmic decision-making», «credit scoring», «machine learning lending», «recommender systems», «explainable AI». Временной охват составил период с 1989 года (публикация модели TAM) по 2025 год. Дополнительно применялся метод «снежного кома»: анализировались списки литературы ключевых работ и публикации, цитирующие их.

Анализ проводился методом тематического синтеза. На первом этапе результаты исследований кодировались по сфере применения, изучаемым детерминантам доверия и поведенческим последствиям. На втором этапе коды объединялись в тематические категории, на основе которых была построена концептуальная модель.

4. Алгоритмическое отторжение и алгоритмическое признание

Исходной точкой современной дискуссии стали эксперименты Б. Дитворста, Дж. Симмонса и К. Мэсси (Dietvorst et al., 2015). Участники, наблюдавшие работу алгоритма и видевшие его ошибки, значительно реже

выбирали его для последующих прогнозов, чем участники, не видевшие ошибок, даже когда алгоритм в целом превосходил человека по точности. Авторы объяснили этот эффект асимметрией ожиданий: от машины ожидается безупречность, поэтому её ошибка интерпретируется как признак системного дефекта, тогда как ошибка человека воспринимается как естественная и простительная.

Систематические обзоры (Burton et al., 2020; Jussupow et al., 2020) выявили широкий спектр причин отторжения: ложные ожидания относительно возможностей алгоритмов, дефицит контроля над решением, отсутствие мотивационных стимулов, несовместимость алгоритмической логики с интуитивными стратегиями принятия решений, а также ощущение, что машина не способна учесть индивидуальные обстоятельства. Показательно, что люди склонны переоценивать своё понимание того, как принимает решения другой человек, по сравнению с тем, как принимает решения алгоритм (Bonezzi et al., 2022): «человеческий чёрный ящик» кажется более прозрачным, чем алгоритмический, хотя на деле это не так.

Вместе с тем Дж. Логг, Дж. Минсон и Д. Мур (Logg et al., 2019) продемонстрировали, что в ряде задач, особенно связанных с числовыми прогнозами, люди охотнее следуют совету алгоритма, чем совету другого человека. Эффект ослаблялся у людей, обладающих экспертизой в соответствующей области, и в ситуации выбора между алгоритмом и собственным суждением.

Противоречие между этими результатами снимается при учёте характера задачи. Н. Кастело, М. Бос и Д. Леманн (Castelo et al., 2019) показали, что готовность полагаться на алгоритм выше в задачах, воспринимаемых как объективные и поддающиеся количественной оценке, и ниже в задачах, воспринимаемых как субъективные, требующие интуиции, вкуса или эмпатии. Важно, что восприятие задачи поддаётся изменению:

подчёркивание количественной природы задачи повышало доверие к алгоритму. Исследование М. Йоманса и соавторов (Yeomans et al., 2019) на примере рекомендаций шуток подтвердило, что алгоритм может превосходить друзей и супругов в прогнозировании предпочтений человека, но люди тем не менее предпочитают рекомендации знакомых, в частности потому, что считают логику алгоритма менее понятной.

Таким образом, отношение потребителей к ИИ не является ни однозначно позитивным, ни однозначно негативным. Оно определяется взаимодействием характеристик задачи, опыта взаимодействия с системой и возможности контроля над её решениями.

5. Концептуальная модель и исследовательские предложения

На основе проведённого синтеза предложена концептуальная модель калиброванного доверия потребителей к ИИ (рисунок 1). Модель рассматривает доверие как результат взаимодействия четырёх групп факторов: контекстуальных, характеристик системы, институциональных и индивидуальных. Воспринимаемая справедливость выступает модератором, усиливающим или ослабляющим влияние этих факторов. Доверие, в свою очередь, определяет поведенческие реакции потребителей, а опыт взаимодействия с системой, включая наблюдаемые ошибки, формирует петлю обратной связи.

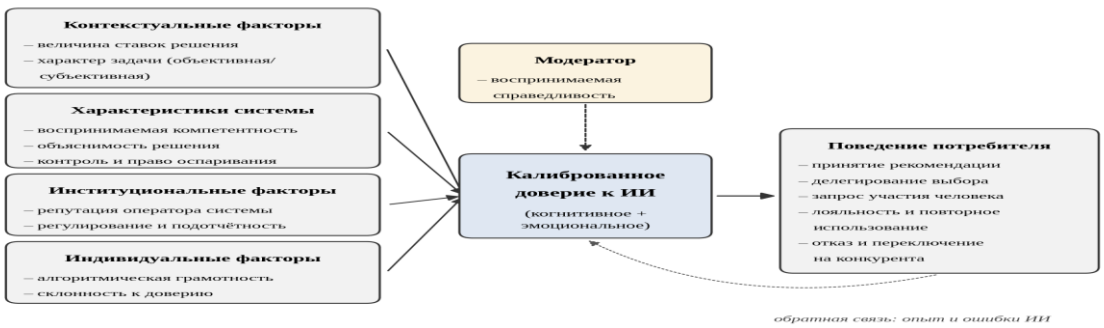


Рисунок 1. Концептуальная модель калиброванного доверия потребителей к ИИ

Источник: разработано автором.

Контекстуальные факторы задают «порог доверия», то есть уровень гарантий, необходимых потребителю для принятия решения ИИ. На основе результатов Кастело и соавторов (Castelo et al., 2019) и Лонгони и соавторов (Longoni et al., 2019) формулируются следующие предложения.

P1. Чем выше воспринимаемые ставки решения, тем выше требования потребителя к объяснимости решения ИИ и возможности участия человека.

P2. Готовность потребителя делегировать решение ИИ выше в задачах, воспринимаемых как объективные и количественные, чем в задачах, воспринимаемых как субъективные и требующие учёта индивидуальных обстоятельств.

Характеристики системы определяют, насколько воспринимаемая надёжность ИИ соответствует реальной. Опираясь на работы Дитворста и соавторов (Dietvorst et al., 2015, 2018), Кизилчека (Kizilcec, 2016) и Шина (Shin, 2021), можно предположить следующее.

P3. Наличие у потребителя возможности скорректировать или оспорить решение ИИ ослабляет эффект алгоритмического отторжения после наблюдения ошибки системы.

P4. Связь между прозрачностью системы и доверием к ней имеет нелинейный характер: объяснения, сфокусированные на причинах конкретного решения и путях его изменения, повышают доверие в большей степени, чем раскрытие общей логики модели.

Институциональные факторы отражают двухуровневую природу доверия к ИИ. В соответствии с моделью Майера и соавторов (Mayer et al., 1995) и её адаптацией к электронной коммерции (McKnight et al., 2002) предлагается следующее.

P5. Доверие к оператору системы и наличие регуляторных гарантий положительно влияют на доверие к ИИ, причём этот эффект сильнее в сферах с высокими ставками, таких как кредитование.

Индивидуальные факторы влияют преимущественно на точность калибровки доверия. Люди, понимающие принципы работы ИИ, в меньшей степени склонны как к слепой вере в алгоритм, так и к его огульному отрицанию (Hoff & Bashir, 2015; Jussupow et al., 2020).

P6. Более высокий уровень алгоритмической грамотности потребителя связан с более точной калибровкой доверия, то есть с меньшим расхождением между воспринимаемой и фактической надёжностью системы.

Наконец, результаты Канемана и соавторов (Kahneman et al., 1986) и исследований алгоритмического кредитования (Bartlett et al., 2022; Fuster et al., 2022) указывают на особую роль справедливости.

P7. Воспринимаемая несправедливость решения ИИ (в частности, различия в ценах или условиях для сопоставимых потребителей) снижает доверие сильнее, чем сопоставимая по экономическим последствиям ошибка прогноза.

6. Обсуждение

Проведённый анализ позволяет ответить на сформулированные исследовательские вопросы. Отвечая на RQ1, можно заключить, что доверие к ИИ формируется на пересечении когнитивной оценки компетентности системы и эмоционального ощущения контроля и безопасности, а его особенностью является двухуровневость: ожидания доброжелательности и честности, неприменимые к алгоритму, переносятся на оператора системы и институциональную среду. Это уточняет классическую модель Майера и соавторов применительно к технологическому посреднику.

Отвечая на RQ2, исследование показывает, что детерминанты доверия существенно различаются в зависимости от сферы. В розничных покупках

доверие формируется через опыт при низких ставках и разрушается прежде всего при обнаружении манипуляции. В кредитовании доверие зависит от объяснимости, справедливости и институциональных гарантий. В сфере услуг ключевую роль играет характер задачи и возможность сочетания ИИ с человеческим участием.

Отвечая на RQ3, следует подчеркнуть, что целью не может являться максимизация доверия к ИИ. Избыточное доверие не менее опасно, чем недостаточное: оно ведёт к автоматизации ошибок, снижению критичности и уязвимости перед манипуляциями. Экономически и социально оптимальным является калиброванное доверие (Lee & See, 2004), при котором его уровень соответствует реальной надёжности системы в данном контексте. Перенос концепции калибровки из инженерной психологии в экономику потребительского поведения представляется одним из значимых теоретических результатов работы.

Кроме того, анализ объясняет кажущееся противоречие между алгоритмическим отторжением и алгоритмическим признанием. Оба эффекта представляют собой проявления одного механизма — оценки соответствия между характером задачи и воспринимаемыми возможностями системы, модерируемой опытом и возможностью контроля.

7. Ограничения и направления будущих исследований

Настоящее исследование имеет ряд ограничений. Во-первых, оно основано на интегративном обзоре литературы, а не на систематическом обзоре с количественным метаанализом, что ограничивает возможность оценки величины эффектов. Во-вторых, значительная часть рассмотренных эмпирических исследований проведена в США и странах Западной Европы, что ставит вопрос о переносимости результатов на иные культурные и институциональные контексты. В-третьих, быстрое распространение

генеративного ИИ опережает накопление эмпирических данных о его влиянии на доверие потребителей.

8. Заключение

Искусственный интеллект не отменяет экономику доверия, а переустраивает её. Если прежде доверие строилось на личных отношениях и репутации институтов, то теперь к ним добавляется новый объект — алгоритм, лишённый собственных намерений, но принимающий решения, значимые для благополучия людей.

Проведённый анализ показывает, что отношение потребителей к ИИ определяется не абстрактной «технофилией» или «технофобией», а конкретными характеристиками ситуации: величиной ставок, характером задачи, объяснимостью, наличием контроля, институциональными гарантиями и воспринимаемой справедливостью. В сфере покупок доверие формируется незаметно, но легко разрушается при обнаружении манипуляции. В кредитовании оно требует объяснимости и регуляторных гарантий. В сфере услуг наиболее устойчивой оказывается гибридная модель, сочетающая возможности ИИ и участие человека.

Предложенная концептуальная модель калиброванного доверия и сформулированные на её основе исследовательские предложения создают основу для дальнейшей эмпирической работы. Конкурентным преимуществом в цифровой экономике становится не столько сам ИИ, сколько способность компании сделать его заслуживающим доверия. В долгосрочной перспективе выигрывают те участники рынка, которые рассматривают доверие не как маркетинговый ресурс, а как обязательство перед потребителем.

Информация об авторе, конфликт интересов и финансирование

Автор заявляет об отсутствии конфликта интересов. Исследование не получало специального финансирования. [При необходимости укажите грант или программу финансирования.]

Список литературы / References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509–514. https://doi.org/10.1126/science.aaa1465

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics, 84*(3), 488–500. https://doi.org/10.2307/1879431

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs, 1*(4), 343–362.

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics, 143*(1), 30–56. https://doi.org/10.1016/j.jfineco.2021.05.047

Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of FinTechs: Credit scoring using digital footprints. *The Review of Financial Studies, 33*(7), 2845–2897. https://doi.org/10.1093/rfs/hhz099

Bonezzi, A., Ostinelli, M., & Melzner, J. (2022). The human black-box: The illusion of understanding human better than algorithmic decision-making. *Journal of Experimental Psychology: General, 151*(9), 2250–2258.

Burton, J. W., Stein, M.-K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 33*(2), 220–239. https://doi.org/10.1002/bdm.2155

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research, 56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668–1674. https://doi.org/10.1126/science.2648573

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science, 64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance, 77*(1), 5–47. https://doi.org/10.1111/jofi.13090

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly, 27*(1), 51–90. https://doi.org/10.2307/30036519