

**УДК 001****Этика исторической памяти человечества в эпоху развития  
искусственного интеллекта****Кидяев Вячеслав Вячеславович**

Аспирант,  
Российский экономический университет им. Г.В. Плеханова,  
главный эксперт Управления по коммуникационной работе  
и международной деятельности  
Национального центра развития искусственного интеллекта  
при Правительстве Российской Федерации,  
109028, Российская Федерация, Москва, Покровский бульвар, 11;  
e-mail: Slavakid@yandex.ru

**Аннотация**

Формирование облика исторической памяти человечества происходило веками, и далеко не всегда этот процесс проходил в условиях справедливости, равенства и этичности. С развитием технического прогресса вопрос создания международного «архива», учитывающего историю стран и народов, этносов и анклавов, встает все более остро. Одной из наиболее подходящих для этого современных технологий, способных обеспечить достаточный уровень автоматизации, безопасности и контролируемости процессов, является искусственный интеллект. Однако на данном этапе технологического развития использование ИИ для создания архива исторической памяти человечества может быть не только недостаточно эффективно, но и небезопасно. Недостаточное использование «мягкого» – этического и «жесткого» – законодательного регулирования этой инновационной сферы, неограниченные полномочия разработчика ИИ-системы, использование ИИ в интересах аффилированных лиц и дискриминация доступа к технологии для отдельных стран сегодня становятся условиями, не только создающими риски, но и сдерживающими развитие доверенных технологий ИИ. Для использования ИИ в международных проектах, таких как создание архива исторической памяти человечества, недостаточно эффективных технологий, они также должны быть полностью регулируемы, подотчетными, общедоступными, а главное – открытыми.

**Для цитирования в научных исследованиях**

Кидяев В.В. Этика исторической памяти человечества в эпоху развития искусственного интеллекта // Контекст и рефлексия: философия о мире и человеке. 2024. Том 13. № 7А. С. 49-56.

**Ключевые слова**

Искусственный интеллект, архив исторической памяти человечества, культурная интерпретация, ангажированность, доверенный ИИ, большие данные, машинное обучение, облачные технологии.

## Введение

Сохранение исторической памяти является неотъемлемым условием для продолжения эволюции человечества. Стоящие «на плечах предков» научные, культурные, географические и иные открытия были бы невозможны без этичного и трепетного отношения к достижениям прошлых поколений. Самым явным историческим подтверждением этого факта выступает даже не технический прогресс человеческого вида, где масштабы развития той или иной технологии напрямую коррелируют с объемом накопленного опыта в этой сфере, а история падения цивилизаций. Регресс и последующая деградация ведущих мировых империй разных периодов всегда были связаны с потерей самоидентификации народов и утратой ими исторической памяти как главного источника аккумуляции знаний внутри этноса/государства/расы. Значительную роль в утрате исторической памяти – информации и опыта, накопленного поколениями, играет дезинформация и намеренное «вымарывание» данных из истории. С развитием цифровых сервисов сохранение исторической памяти стало играть еще более важную роль, так как риск замещения реальных данных ложными вырос параллельно с эволюцией технологий. Однако критическую массу эта проблема приобрела с появлением технологий искусственного интеллекта (ИИ).

### Потенциал внедрения искусственного интеллекта в чувствительные отрасли экономики

Популяризация технологий машинного обучения и ИИ позволили значительно оптимизировать и автоматизировать ключевые отрасли экономики и социальной сферы стран. Роботизация производств, внедрение систем видеонаблюдения на основе технологии компьютерного зрения, создание цифровых двойников основных производственных узлов или отдельных частей производства позволили прогнозировать наступление аварий и чрезвычайных ситуаций, снизить число брака в выпускаемой продукции и обеспечить беспрецедентный уровень безопасности. Однако технология генеративного искусственного интеллекта, развивающаяся быстрее других ИИ-технологий, помимо позитивных эффектов, создала ряд рисков, в первую очередь связанных с центральными этическими дилеммами человечества [Кодекс этики в сфере искусственного интеллекта, [www](#)].

Вопросы о цифровой реинкарнации, этике использования медицинских ИИ-решений при диагностике заболеваний, защите частной жизни и в особенности распространения ложной информации перешли из категории философских в практическую плоскость. Очевидно, что совершенствование генеративного ИИ создало как в мировой сети, так и вне ее прецедент, когда появление новых данных перестало быть ограниченным чем-либо. Однако созданные с помощью ИИ данные не являются отражением реальности или ее «продуктом». Появившиеся таким образом данные – искусственные, так как не являются напрямую результатом человеческого труда, а генерируются алгоритмами. Такая информация не считается сама по себе ложной, потому что генерируется на основе реальных библиотек данных, но сам инструмент генеративного ИИ несет в себе устрашающий, неограниченный потенциал для создания порочащих и неэтичных сведений, новостей и материалов. Это провоцирует множество неконтролируемых рисков, связанных с сохранением культурной и исторической памяти отдельных государств и даже всего человечества.

Технологии ИИ уже изменили подход человечества к сохранению и интерпретации

---

исторической памяти. Сейчас значительная часть работы с архивированием, сохранением и анализом прошлого обеспечивается алгоритмами ввиду устройства самой технологии. В обработке большого объема однотипных данных ИИ на порядок превзошел все достижения человеческого «ручного» труда. Но с расширением его возможностей возникает логичный этический вопрос: может ли ИИ захватить контроль над исторической памятью человечества, если не будет вовремя ограничен нормами этики? [Белая книга искусственного интеллекта: европейский подход к совершенству и доверию, www]. Этот вопрос выходит за рамки технического устройства ИИ и нейросетей и требует глубокого философского анализа и человеческой интерпретации, юридического и морального осмысления последствий неконтролируемого и неравномерного развития технической и этической составляющих ИИ.

Одной из особенностей исторической памяти является ее культурная интерпретация. В отличие от генерации данных алгоритмами ИИ, человеческое общество формировало историческую память непоследовательно и неоднородно, часто опуская негативные ее части и гиперболизируя успехи. Это субъективный процесс, в центре которого, кроме объективных фактов, находится также человек с его убеждениями, моральными границами и системой ценностей. Философы Морис Хальбвак и Пол Рикёр утверждали, что память включает в себя как коллективные, так и индивидуальные элементы, может быть селективной, интерпретативной или политически ангажированной. Подобная классификация исключает субъективность при накоплении исторической памяти, но обеспечивает определенный уровень нейтральности. Пересечение неоднородных взглядов разных объединений (государственных, этнических, политических и иных) позволило человечеству создать некий «стандарт» в понимании общего прошлого. Безусловно, выработанный шаблон невозможно назвать сколь угодно справедливым и уж тем более объективным, но он по самой своей сути является производным человеческого разума и накопленного человечеством опыта – исторической памятью.

### **Кризис доверия при использовании искусственного интеллекта для управления архивами**

Допуск ИИ к управлению данными об истории человечества в краткосрочной перспективе может принести значительную пользу. Обработка и систематизация данных с помощью ИИ позволяет в сотни раз ускорить процессы архивирования и классификации исторических событий и сведений о прошлом. Создание специализированных баз данных уже сейчас открыло перед человечеством такую перспективу для определения самого себя, какой не давала ни одна из ранее созданных технологий. Вместе с тем, отдавая в ведение ИИ управление исторической памятью и позиционируя алгоритмы как универсальный и непредвзятый способ трактовки событий, человечество добровольно отказывается от управления ей. Но действительно ли ИИ и нейросети способны выступать как независимая, объективная сторона, лишенная человеческих слабостей и ограничений?

Однозначного ответа на этот вопрос не существует. Однако на данном этапе развития технологий очевиден факт того, что любая нейросеть или ИИ-система будет подчинена тем правилам, условиям и принципам, которые заложены в нее разработчиком на этапе создания. Именно команда разработчиков определяет моральные границы и формирует позицию нейросети по ключевым вопросам. Также очевидно, что любой разработчик будет программировать ИИ в соответствии со своими убеждениями или убеждениями заказчика.

Отсутствие у ИИ возможности для самовоспроизведения на данном этапе эволюции технологии исключает возможность создания действительно независимой, непредвзятой и лишенной предрассудков нейросети. Это доказывает результат исследований, проведенных в системе здравоохранения Москвы в 2023 году. Сервисы медицинского ИИ, призванные помогать врачу при диагностике рака кожи – меланомы, в несколько раз хуже считывали данные жителей Москвы, приехавших из стран Центральной Азии и Африки. Это произошло из-за того, что база данных, на которой был обучен ИИ, обладала недостаточным количеством токенов информации по типам кожи, отличным от европеоидной. Таким образом, создание и эксплуатация ИИ для сохранения человеческой памяти в данный момент не может быть универсальным и даже безопасным способом обращения с ней. Доверие к ИИ-алгоритмам может стать ловушкой для человечества, загружающего огромные объемы данных на облачные сервера под управлением ИИ, так как это снижает степень контроля над процессами накопления и структурирования исторической памяти.

Еще одной крупной проблемой, частично затронутой в начале данной статьи, является проблема интерпретации. Ключевым условием сохранения исторической памяти человечества является наличие «человеческого фактора» в этом процессе. Действительно, это может создавать ряд рисков искажения памяти в интересах отдельных групп, но это открывает и значительные возможности для четкого и последовательного выстраивания человечеством собственного морального контура, самоопределения им границ допустимого зла, а также вычленения благодетелей и их признания. Это значит, что даже если исторически сила всегда тождественна насилию, а отказ от насилия в большинстве сценариев приравнивается к незащитности и слабости, историческая память человечества не допустит слепого культа силы и попраiania благодетели, руководствуясь исключительно измеримой эффективностью обоих методов [Мухаев., 2022]. Сегодня генеративный ИИ демонстрирует также значительные проблемы с появлением «галлюцинаций» – особых искажений в генерации, результаты которых полностью или почти полностью бессмысленны, а причины появления неизвестны. Галлюцинациям подвержены даже большие языковые модели, созданные самыми передовыми технологическими компаниями, что делает выбор ИИ как единственного инструмента для сохранения и накопления исторической памяти не только опасным, но и неоправданно рискованным. При этом считать, что человечество откажется от ИИ в работе с большими данными и архивами, бессмысленно. Этот инструмент позволяет значительно оптимизировать процессы обработки данных, а аналогичных по эффективности безопасных и доверенных разработок сегодня просто нет.

Полная передача работы с большими данными в ведение ИИ не кажется фантастикой сегодня. Но именно контролируемые ИИ органы или структуры будут способны в полной мере распорядиться им. Так возникает проблема свободы доступа ко всеобщему достоянию человечества – исторической памяти. Группа, контролирующая обработку данных, получит доступ к самостоятельной интерпретации мировой истории [«Искусственный интеллект в контексте культурного наследия и музеев: сложные задачи и новые возможности», [www](http://www)]. Это создает риск искажения воспоминаний в интересах конкретных лиц, что представляет значительную опасность для сохранения истинного облика человеческой истории и культуры. Также риски могут возникнуть непреднамеренно, при недостаточном уровне контроля за деятельностью нейросетей. Например, допущение ошибок или генерация ложных данных могут значительно повлиять на сохранение исторической памяти, если они допущены ИИ в важных архивах и базах данных, аналогичных тому, о котором идет речь в данной статье. Как ранее

---

было отмечено, на текущем этапе развития технологий создать 100% доверенный ИИ невозможно, а ручная проверка результатов деятельности технологии в сути своей делает бессмысленным весь организованный с помощью ИИ процесс автоматизации. Парадоксальность ситуации заключается в том, что определить действительно значимые искажения в архивах своевременно почти невозможно. Однако они станут очевидными сразу, как только ситуация выйдет из-под контроля и ее уже нельзя будет решить быстро и без последствий.

### **Международная этика и инновационные технологии**

Одним из эффективных решений, позволяющих еще на этапе создания ИИ снизить потенциальные риски его влияния на историческую память, является междисциплинарный подход к разработке. Привлечение к созданию новой модели ИИ философов, антропологов, социологов, историков и специалистов по этике позволит в значительной степени сократить потенциальные риски при дальнейшей ее эксплуатации. Кроме того, это обеспечит лучшее понимание того, как люди смотрят на историческую память, что вкладывают в это понятие и каким видят истинно моральное обращение с ней. Это даст возможность человечеству создать относительно прозрачные алгоритмы, подчиненные не конкретной организации или стране, а цели, которая закладывалась в них на этапе разработки. Возможность открытия доступа к данным и алгоритмам их обработки позволит обезопасить ИИ от скрытых манипуляций и внутреннего лоббизма. Это реализует права людей на доступ к информации, прописанные, кроме прочего, в одном из федеральных законов Российской Федерации [Федеральный закон №149 от 27 июля 2006 года, [www](http://www)]. В идеальных условиях это даст всем членам общества в равной степени поучаствовать в формировании исторической памяти человечества, а также в ее дальнейшей актуализации.

Моральная ответственность за сохранение правдивой и справедливой исторической памяти лежит на всем человечестве. И хотя искажение памяти с помощью, например, маргинализации отдельных этносов или групп было неотъемлемой частью формирования исторической памяти человечества, это не значит, что с этим не нужно бороться. И напротив, в человеческой истории есть множество примеров, когда отдельная нация или объединение совершало ряд поступков, полностью противоречащих общечеловеческим нормам морали, что приводило к намеренному «вымарыванию» любых положительных упоминаний из исторической памяти. Лучшим примером служит фашистская Германия середины XX века, которая после поражения во Второй мировой войне лишилась не только своей человеконенавистнической идеологии, но и права на сколь угодно положительную или даже нейтральную оценку политического и социального строя того времени. Решения о намеренном «изгнании» отдельных периодов истории некоторых стран порой делается извне, как в случае с Германией, когда страну принудили к осуждению своего прошлого, так и намеренно, изнутри. Во втором случае примером могут служить Соединенные Штаты Америки, руководящие органы которых намеренно и на протяжении нескольких веков занижали масштабы геноцида коренного населения Северной Америки – индейцев.

Подобные этические «нюансы» при формировании истории общего человеческого прошлого доказывают, что неизбирательное использование даже полностью правдивых событий и отдельных фактов не может являться универсальным подходом из-за особенностей структуры человеческого общества и его моральных ориентиров. Создать доверенную ИИ-

систему, способную обеспечить архивирование и хранение данных о человеческой истории в том виде и объеме, в котором это необходимо человечеству сегодня, – возможно. Намного более сложными задачами являются организация разработки такой системы на базе международных институтов, создание контролирующих органов, деятельность которых в достаточной степени удовлетворит мировое большинство, а также определение уровня «открытости» потенциального общечеловеческого архива. Решение этих вызовов требует не только и не столько определенных достижений технологического прогресса, сколько прогресса в области социологии, международной дипломатии, этики ИИ, а главное – философии. Именно межотраслевое международное регулирование может стать гарантией того, что потенциальная разработка не только будет отвечать всем технологическим требованиям человечества, но и обеспечит справедливый доступ стран к общей исторической памяти, а также подотчетность разработчиков системы мировому сообществу и полную прозрачность в выборе инструментов для создания архива.

### **Заключение**

Таким образом, вопросы, связанные с переносом данных архивов, касающихся исторической памяти отдельных стран и всего человечества в целом, не должны решаться без должного международного и межотраслевого предварительного анализа. Желание разработать инструмент для накопления исторической памяти человечества «в цифре» и под управлением ИИ-технологий создает ряд неконтролируемых рисков, среди которых риски доступа к информации, риски контроля за созданием и эксплуатацией технологии, риски неэтичного формирования базы данных, а также риски недостаточного контроля за ИИ и невозможность внесения достаточных правок вручную при возникновении подобной необходимости и так далее. Перечень рисков при создании архива исторической памяти человечества под управлением ИИ настолько широкий, что единственной возможностью их преодоления на данном этапе видится лишь формирование международной экспертной команды, состоящей из межотраслевых специалистов и имеющих разный опыт, возраст, пол и национальную принадлежность для того, чтобы обеспечить удовлетворенность как-можно большего количества людей. Еще одной проблемой, решение которой лежит уже в плоскости технологического прогресса, является невозможность создания полностью доверенной и безопасной ИИ-технологии, которой можно было бы вверить управление подобным мегалитическим архивом данных. Параллельное решение вопросов как этического, так и технического характера с достаточным уровнем международной открытости при разработке позволят не только приблизить появление архива исторической памяти, но и гарантируют сохранение исторической памяти в наиболее полном виде и объективной форме, что не было доступно ни одному поколению людей до сегодняшнего дня.

### **Библиография**

1. Альянс в сфере искусственного интеллекта. «Кодекс этики в сфере искусственного интеллекта». URL: [https://ethics.a-ai.ru/assets/ethics\\_files/2023/05/12/Кодекс\\_этики\\_20\\_10\\_1.pdf](https://ethics.a-ai.ru/assets/ethics_files/2023/05/12/Кодекс_этики_20_10_1.pdf) (дата обращения 19.09.2024).
  2. Белая книга искусственного интеллекта: европейский подход к совершенству и доверию. URL: [https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en](https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en) (дата обращения: 20.09.2024).
  3. Мужаев Р.Т. Медиаполитика. М.: ИНФРА-М., 2022. 401 с..
  4. Искусственный интеллект в контексте культурного наследия и музеев: сложные задачи и новые возможности.
-

URL: [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2023\)747120](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2023)747120) (дата обращения: 25.09.2024).

5. Об информации, информационных технологиях и о защите информации: федер. закон от 27 июля 2006 года № 149. URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_61798](https://www.consultant.ru/document/cons_doc_LAW_61798) (дата обращения: 28.09.2024).

## **Ethics of the historical memory of humanity in the era of artificial intelligence development**

**Vyacheslav V. Kidyaev**

Postgraduate Student,  
Plekhanov Russian University of Economics,  
Chief Expert of the Department of Communications and International Affairs  
of the National Center for Artificial Intelligence Development  
under the Government of the Russian Federation,  
109028, 11 Pokrovsky Boulevard, Moscow, Russian Federation;  
e-mail: Slavakid@yandex.ru

### **Abstract**

The historical memory of mankind is a system of historical messages transmitted from generation to generation, myths about the events of the past, subjectively comprehended and reflected by generations of people, individual states and ethnic groups in their own interests. The loss of historical memory can become one of the main reasons that can lead to both the decline of the human race and its complete extinction. Huge efforts are being made to preserve it today: megalithic archives of historical data are being created, and states are striving to restore reliable events of past historical periods. Artificial intelligence has become the game changer technology. Thanks to AI's ability to work with big data, humanity's ability to preserve historical memory has increased hundreds of times. But along with new opportunities, new threats have emerged. Due to the frankly weak security of all the world's AI developments today, as well as its mass introduction, humanity is closer than ever to the moment when our historical memory may be distorted or even completely lost. In such circumstances, the inclusion of international intergovernmental institutions in the development of projects of planetary importance, based on AI solutions, may be the only way to guarantee the safety and long-term effectiveness of such megaprojects.

### **For citation**

Kidyaev V.V. (2024) *Etika istoricheskoi pamyati chelovechestva v epokhu razvitiya iskusstvennogo intellekta* [Ethics of the historical memory of humanity in the era of artificial intelligence development]. *Kontekst i refleksiya: filosofiya o mire i cheloveke* [Context and Reflection: Philosophy of the World and Human Being], 13 (7A), pp. 49-56.

### **Keywords**

Artificial intelligence, archive of historical memory of mankind, cultural interpretation, engagement, trusted AI, big data, machine learning, cloud technologies.

## References

1. Alliance in the field of Artificial Intelligence. "Code of Ethics in the field of Artificial Intelligence". URL: [https://ethics.a-ai.ru/assets/ethics\\_files/2023/05/12/Кодекс\\_этики\\_20\\_10\\_1.pdf](https://ethics.a-ai.ru/assets/ethics_files/2023/05/12/Кодекс_этики_20_10_1.pdf) (date of access 09/19/2024).
  2. Artificial Intelligence in the Context of Cultural Heritage and Museums: Complex Challenges and New Opportunities. URL: [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2023\)747120](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2023)747120) (date of access: 25.09.2024).
  3. Mukhaev R.T. Media Policy. Moscow: INFRA-M, 2022. 401 p..
  4. On information, information technology and information protection: federal law of 27 July 2006 No. 149. URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_61798](https://www.consultant.ru/document/cons_doc_LAW_61798) (date of access: 28.09.2024).
  5. White Paper on Artificial Intelligence: A European Approach to Excellence and Trust. URL: [https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en](https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en) (date of access: 09/20/2024).
-