**Мешков Владимир Евгеньевич,**
кандидат технических наук, доцент,
независимый исследователь
Meshkov Vladimir Evgenevich,
independent researcher

**Чураков Вадим Сергеевич,**
кандидат философских наук,
независимый исследователь
Churakov Vadim Sergeevich,
independent researcher

# ПОТЕНЦИАЛЬНЫЕ РИСКИ РЕТРОСПЕКТИВНОГО ВМЕШАТЕЛЬСТВА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В БАЗЫ ДАННЫХ
# POTENTIAL RISKS OF RETROSPECTIVE INTERFERENCE BY ARTIFICIAL INTELLIGENCE IN DATABASES

**Аннотация.** В статье рассматриваются системные риски, возникающие при использовании технологий искусственного интеллекта (ИИ) для ретроспективного анализа, коррекции и модификации исторических данных в автоматизированных базах данных (БД). Процессы автономной актуализации архивных массивов ИИ-системами создают новые уязвимости для целостности и достоверности информации.

**Abstract.** The article examines the systemic risks that arise when using artificial intelligence (AI) technologies for retrospective analysis, correction, and modification of historical data in automated databases (DBs). The processes of autonomous updating of archival data sets by AI systems create new vulnerabilities to the integrity and reliability of information.

**Ключевые слова:** Нейросети, искусственный интеллект, виртуальная реальность, риски, ретроспективное вмешательство, сетевая модель времени.

**Keywords:** Neural networks, artificial intelligence, virtual reality, risks, retrospective intervention, network time model.

В данной работе авторы рассматривают риски от несанкционированного ретроспективного вмешательства в базы данных.

Время существует в трёх ипостасях: прошлое, настоящее и будущее. Понятие и представление времени в искусственных системах обладает своей спецификой и рядом особенностей. Данная статья является продолжением работ авторов по тематике времени в искусственных системах [1].

В действующих искусственных системах образуются свои особые темпоральность и виртуальная реальность. Виртуальная реальность и искусственный интеллект – две революционные технологии, которые в последние годы набирают огромную популярность и влияют на множество сфер нашей жизни. Их взаимодействие открывает новые горизонты для создания инновационных решений и уникальных пользовательских опытов.

Совмещение виртуальной реальности с искусственным интеллектом позволяет создать интерактивные и адаптивные среды, где пользователи могут не только наблюдать, но и взаимодействовать с цифровыми мирами. Современные нейросети, чат боты и алгоритмы, обрабатывающие петабайты данных, оперируют временными метками и последовательностями, но не испытывают «длительности» в бергсоновском смысле. Для них прошлое, настоящее и будущее - не переживаемый поток, а структурированные массивы

информации, доступные для анализа вне контекста темпоральности. Естественный интеллект/ интеллект человека – воспринимает время как объективную реальность, и ему присуща темпоральная субъективность, которая накладывает ограничения как на изучение времени, так и на возможности оперирования временем, а в интеллекте искусственном данная темпоральная субъективность отсутствует, что и позволяет ему оперировать временем не как фундаментальной сущностью, а как инструментальной метрикой.

Кроме того, нейросети обрабатывают информацию пакетами (батчами) или кадрами, разбивая непрерывный поток на отдельные состояния (т.е. самым производя дискретизацию процессов). А также им присуща эмерджентность паттернов. Временные зависимости в данных (например, тренды в соцсетях) – не свойство «времени», а результат корреляций между точками.

При этом, существует проблема многовариантности начальной модели (начальной установки) нейросети. Следовательно, могут быть созданы различные варианты образов в нейросети.

Еще одним аспектом, на который следует обратить внимание, является то, что у нейросети нет «личного прошлого» в человеческом понимании (у неё нет воспоминаний, опыта жизни или сознания). Это может повлиять на организацию процесса ретроспективного анализа. Однако, у нейросети есть «история создания и развития» как технологии, а также «прошлое» в виде данных, на которых она обучалась. У сети есть «прошлое» в виде обучающей выборки: нейросеть изучает тексты, изображения и другие данные за определённый период. Следовательно, есть возможность ретроспективного вмешательства в данные за этот период.

Виртуальная реальность вопреки современному мировидению, в котором существует чёткое разграничение между прошлым, настоящим и будущим, где линейное восприятие времени, сопряжено с идеей необратимости, вновь обращает нас к мифологическому восприятия времени. Виртуальная реальность возвращает нас к целостному восприятию мира, где прошлое и будущее имплицитно присутствуют в настоящем [1].

Возможное отсутствие времени в виртуальных пространствах приводит и к ещё одному значимому онтологическому эффекту – нарушаются причинно-следственные отношения. В реальности каузальность предполагает необходимое следование одного события за другим, а в виртуальных пространствах традиционное расположение событий в отношении «прошлое-будущее» может и не наблюдаться.

Следовательно, «в причинно-следственных связях можно обнаружить не тривиальные временные зависимости одного события от другого, не простое их следование друг за другом, а их сущностные отношения» [2].

Отметим, что сетевая модель времени применима для манипуляций виртуальной реальностью, посредством ретроспективного изменения различных баз данных.

Рассмотрим некоторые варианты ретроспективного вмешательства в базы данных, для чего может быть использована программа «виртуальный терминатор». Программа агент виртуальный терминатор отравляется в прошлое, чтобы откорректировать базы данных.

Отметим, что «виртуальный терминатор» может представлять собой мультиагентную систему (совокупность агентов, действующих автономно для достижения поставленной цели).

Виртуальный терминатор может также работать с региональной или глобальной нейронной сетью, которую мы рассматриваем как некоторую память. Так как, память – «это различимый наблюдателем стабильный след от внешнего воздействия», то целенаправленное воздействие изменяет память, следовательно, изменяет представление об объекте [3].

Поиск в сети объектов для модификации (потенциальной модификации) осуществляется поэтапно (по тактам виртуального времени). Входом в сеть (началом алгоритма) будем считать узел базового модифицируемого понятия (ситуация, состояние и

т.д. в зависимости типа нейронной сети). Уточним, что мы понимаем под прошлым информационного объекта. В данном случае возможны два варианта описания объекта:

1) Изменение параметров объекта во времени с ретроспективным сохранением значений, например, в темпоральных базах данных [4].

2) Существование объекта во времени многоплановой виртуальной реальности.

Рассмотрим в данной работе первый вариант. Пусть мы имеем возможность доступа к описанию информационного объекта. Тогда требуется изменить не только конкретное описание объекта, но и описание образа этого объекта в других объектах. Будем считать переход к другим объектам переходом во времени назад. Хотя, пока нашего объекта нет, не может быть и его образа в сети. Отметим, что «прошлое» у объекта возникает только тогда, когда объект модифицируется. Это же можно сказать и об образе объекта в других объектах.

Если наш объект скрыть в виртуальном пространстве, а видимыми сделать его образы в других объектах, то можно продвинуть новое (другое) представление о нашем объекте. Сторонние образы объекта не являются его прошлым (объект создан раньше), но могут быть его настоящим и будущим. Однако, если события или объекты изменяются динамично, но предыдущие значения сохраняются, то тогда у них формируется прошлое, которым и можно манипулировать.

Определим, что «ретроспективное вмешательство» это сдвиг во времени назад. Временной сдвиг или сдвиг во времени – это разница между первоначальным и последующим перемещением объекта на ременной шкале (влево или вправо в прошлое или в будущее) с определённым шагом.

Рассмотрим возможные варианты (цели) вмешательства:

- «обнуление» (уничтожение объекта в виртуальной реальности);
- «легализация» (создание объекта в виртуальной реальности);
- «модификация» объекта с некоторой целью (положительной или отрицательной с какой-то точки зрения).

При этом существуют риски отклонения от поставленных целей.

Если представить виртуальное пространство нейросетевой моделью, то на каждом шаге ретроспективного погружения мы будем иметь свой вариант модели [5]. В этом случае возникает проблема определения временного интервала для формирования ретроспективной модели. Выбор дискрета квантования по времени либо линеен, либо не линеен. В этом случае, прежде всего, необходимо определиться с глубиной погружения в прошлое. Можно рассматривать два варианта:

- погружение до тех пор, пока находится информация об объекте исследования;
- погружение на заранее заданную глубину.

Отметим, что предлагаемые авторами методики в основном применимы в специализированных нейронных сетях, позволяющих вносить соответствующие изменения. С учетом идей, высказанных в [6], возможно хранить фрагменты нейросети до и после воздействия. Тогда появляется возможность многовариантности и восстановления исходной версии. Можно заранее сгенерировать целый ряд виртуальных образов для объекта, и использовать их по мере необходимости.

Не будем лукавить и отметим, что речь идёт не о хакерах, а о владельцах конкретных нейронных сетей. Причем основные риски связаны с потенциальным несанкционированным вмешательством ИИ в информационные поля нейронной сети.

*Список литературы:*

1. Мешков В.Е., Мешкова Е.В., Кравченко П.Д., Чураков В.С. Представление времени в искусственных системах: монография /Под ред. В.С. Чуракова (Серия «Библиотека времени». Вып.15). - Ростов на-Дону- Новочеркасск: Изд-во ‹‹НОК››, 2019.

2. Пилипенко Е. А. Виртуальное время: онтологический аспект // Грамота.2015. №12 (62): в 4-х ч. Ч.1.

3. Денисенко А.В. Асинхронный подход к моделированию нейронной сети. Часть I. Время как нематериальная среда памяти//Проблемы моделирования нейронных сетей. 2018. № 19(1).

4. Кольчугина Е.А. Искусственные миры: темпоральная организация//Программные продукты и системы. № 2. 2013.

5. Осипов В.Ю. Нейронная сеть с прошедшим, настоящим и будущим временем//Информационно-управляющие системы. №4.2011. // cyberleninka.ru›article/n/neyronnaya-set-s- …

6. ИИ с двойным дном//colonelcassad.livejournal.com›10257372.html.