Возможности нейротехнологий в оценке потенциально противоправного аудиовизуального контента: обоснование гипотезы

Рашитханов Руслан Салихович
заместитель директора, Институт информационной и медиабезопасности, Университета имени О.Е. Кутафина (МГЮА), RSRASHITHANOV@msal.ru

В статье обосновывается научная состоятельность гипотезы о возможности использования нейрофизиологических показателей для объективной оценки негативного влияния аудиовизуального контента на психофизиологическое состояние человека и, соответственно, принятия правовых мер ограничения оборота такого контента. Рассматриваются теоретико-правовые и этические ограничения применения нейрофизиологических данных в контексте регулирования цифрового контента. Обосновывается вспомогательная роль нейротехнологий как источника дополнительной информации о риске контента, дополняющего традиционные психологические, лингвистические и правовые экспертные подходы, но не подменяющего их. Описана концепция эмпирической верификации научной гипотезы: исследование создает основу для реализации правовых механизмов противодействия деструктивному контенту и установления того, каким образом результаты нейропсихофизиологических измерений могут быть интегрированы в процедуры экспертной оценки контента при сохранении приоритета прав и свобод человека. В рамках данной статьи обосновывается состоятельность гипотезы и предлагается программа эмпирического исследования, необходимого для её проверки. При этом вопросы, касающиеся проведения экспериментов, обработки данных и верификации результатов выходят за рамки данной работы и составляют предмет отдельного исследования.

Ключевые слова: нейротехнологии, нейрофизиологические показатели, ЭЭГ, аудиовизуальный контент, противоправный контент, нейроправо, когнитивный суверенитет.

В современном информационном обществе проблема охраны психического здоровья личности приобретает особую значимость в связи с лавинообразным ростом контента и усложнением механизмов его воздействия на сознание человека. Аудиовизуальные материалы, распространяемые через средства коммуникации, способны оказывать как позитивное, так и деструктивное влияние на психическое состояние индивида. В последние десятилетия юридическая наука и практика столкнулись с необходимостью выработки объективных критериев оценки вреда, причиняемого информационной продукцией психическому здоровью граждан.

Предлагаемая модель исследования опирается на быстро развивающееся направление исследований, в которых нейрофизиологические показатели человека регистрируются в ответ на предъявление аудиовизуального контента. В ряде работ, прежде всего в области нейромаркетинга и компьютерного распознавания эмоций (аффективных вычислений), показано, что комплексное использование ЭЭГ, показателей сердечно-сосудистой системы и кожно-гальванической реакции (КГР/EDA) позволяет объективно оценивать уровень эмоционального возбуждения, вовлеченности при просмотре видеороликов, рекламных материалов и других динамических стимулов. Электродермальная активность, как отмечают некоторые ученые, является особенно чувствительным и относительно «чистым» маркером симпатической активации, тесно связанной с эмоциональным возбуждением и вниманием, что делает ее удобным объективным индикатором эмоционального состояния при воздействии медийных стимулов. Также исследования демонстрируют возможность автоматического распознавания эмоциональных состояний по EDA, в том числе с использованием фрактального анализа и современных методов машинного обучения, а также стабильное распознавание эмоций по ЭЭГ-сигналу с применением глубоких нейронных сетей. Всё это свидетельствует о том, что эмоциональная и стрессовая реакция на сложный аудиовизуальный контент может быть описана в терминах объективных физиологических метрик, а не только субъективных ощущениях и оценках. Проще говоря, возможность объективной фиксации состояния человека посредством упомянутых технологий прослеживается явно.

Вместе с тем принципиально важно разграничивать регистрацию физиологической реакции и ее содержательную интерпретацию. Повышение электродермальной активности, изменение сердечного ритма, вариации ЭЭГ-ритмов или амплитуды событийно-связанных потенциалов свидетельствуют о наличии определенной реакции организма, но сами по себе не позволяют однозначно установить ее психологический смысл. Один и тот же физиологический паттерн может быть связан с тревогой, интересом, удивлением, возбуждением, когнитивным усилием или реакцией на новизну. Поэтому в настоящей статье нейрофизиологические показатели рассматриваются как измеряемые признаки, требующие интерпретации в совокупности с самоотчетами, поведенческими данными, характеристиками стимула и контекстом его предъявления.

Особый интерес для настоящего исследования представляют воздействия насильственного и в целом «стрессогенного» медиаконтента на человека. Исследования с использованием МРТ и ЭЭГ показали, что при просмотре

Статья подготовлена в рамках программы "Приоритет 2030"

сцен насилия в телепередачах, фильмах и видеоиграх активируются структуры, ответственные за обработку угрозы, внимания и эмоционального возбуждения (орбитофронтальные и префронтальные отделы коры головного мозга), а характер и выраженность этих активаций зависят от истории предыдущего воздействия. В одной из научных работ по рассматриваемой проблематике явно показано, что длительное увлечение насильственными видеоиграми связано со снижением амплитуды компонента P300 при предъявлении изображений насилия, причем степень снижения P300 предсказывала уровень агрессивного поведения в последующей задаче. Другие исследования эмоциональной перцепции и задач на торможение реакции показывали изменение фронтальных компонентов N200/P300 после просмотра насильственных фильмов, что интерпретируется как нейрофизиологическая десенсибилизация к насилию и изменение когнитивного контроля. Вместе с тем имеются и работы, не подтверждающие гипотезу о простом и линейном эффекте десенсибилизации, что подчёркивает сложность и многомерность взаимосвязей между типом контента, нейрофизиологическими ответами и поведенческими последствиями.

На этом фоне целесообразно сформулировать гипотезу: **нейрофизиологические показатели (ЭЭГ, фМРТ или fNIRS, ЭКГ, КГР/EDA, показатели дыхания и др.) при предъявлении невербального аудиовизуального контента (включая психоделические, интенсивно воздействующие визуально-аудиальные ряды) позволяют количественно описать меру негативного влияния этого контента на психофизиологическое состояние человека**. Под «негативным влиянием» в рамках исследования предполагается понимать операционализируемые характеристики: уровень стрессовой активации, выраженность отрицательной эмоциональной реакции, когнитивную перегрузку, устойчивость реакции во времени, а при наличии лонгитюдной части – длительности последствий. Исходя из существующих данных о том, что по физиологическим каналам возможно надежно классифицировать эмоциональные состояния, в том числе при предъявлении динамических видео и музыкальных стимулов, предполагается, что для особого класса контента можно выделить характерные паттерны нейрофизиологической реакции, отличающиеся от реакций на нейтральные и контролируемые эмоциональные стимулы.

В рамках данной статьи обосновывается состоятельность гипотезы и предлагается программа эмпирического исследования, необходимого для её проверки. При этом вопросы, касающиеся проведения экспериментов, обработки данных и верификации результатов выходят за рамки данной работы и составляют предмет отдельного исследования.

В рамках предполагаемого исследования, необходимость которого обосновывается в данной статье, цель состоит в разработке и валидации научно обоснованной методики объективной оценки психофизиологического влияния невербального цифрового контента на человека на основе комплексной регистрации нейрофизиологических сигналов. Для достижения этой цели потребуется решить несколько взаимосвязанных задач.

Во-первых, необходимо сформировать репрезентативный набор стимулов, включающий нейтральный контент, стандартизированные эмоциональные видеоклипы, используемые в современной психофизиологии (например, базы типа MAHNOB-HCI, в которой собраны видео- и аудиоданные, а также реакции людей на них), в том числе контент, вызывающий сомнения (психоделические, потенциально деструктивные видеоряды и др.).

Во-вторых, требуется разработать экспериментальные протоколы, обеспечивающие одновременную регистрацию многоканальной ЭЭГ, EDA, ЭКГ, дыхания и, по возможности, методов нейровизуализации (фМРТ или, более реалистично, fNIRS) при предъявлении указанных стимулов в стандартизированных условиях.

В-третьих, предстоит сформировать систему количественных показателей (индексов) эмоционального возбуждения, валентности, внимания и когнитивной нагрузки на основе современных подходов к анализу соответствующих сигналов, включая фрактальные и спектральные методы для EDA, методы выделения событийных потенциалов из ЭЭГ, а также применение машинного обучения и глубоких нейросетевых моделей для построения классификаторов и регрессионных моделей эмоционального состояния.

Новизна предлагаемого исследования заключается, во-первых, в переносе хорошо разработанных нейрофизиологических методик на специфический класс контента, который трудно поддается традиционной экспертизе – невербальный контент, аморфные или психоделические аудиовизуальные стимулы. Если в существующих исследованиях основное внимание уделяется либо рекламным материалам, либо явно насильственному контенту (фильмы, видеоигры), то в рамках планируемого исследования предполагается систематически изучать именно те случаи, где вербальные признаки отсутствуют или минимальны, а воздействие осуществляется через форму, цвет, динамику, ритм и звуковую структуру.

Во-вторых, новизна состоит в разработке интегральных индексов «нейрофизиологического риска» контента, которые опираются не на один канал (например, только P300 или только EDA), а на сочетание нескольких сигналов, что повысит устойчивость и обобщаемость оценок. На практике мультиканальные подходы уже показали преимущества при оценке реакции на медийные стимулы: совмещение ЭЭГ, КГР и сердечного ритма позволяет лучше предсказывать субъективные оценки, чем каждый канал по отдельности.

В программе будущего исследования необходимо реализовать сравнительный подход с участием нескольких групп испытуемых, различающихся по опыту потребления медиаконтента и уязвимости (например, молодые пользователи с высоким временем ежедневного онлайн-просмотра, клинические или субклинические группы с повышенной тревожностью, контрольная группа с умеренным медиапотреблением). Планируется, что при предъявлении различных типов стимулов будет проводиться анализ как усредненных по группе паттернов, так и вариативности. Особое внимание должно быть уделено разделению компонентов реакции, связанных с новизной и удивлением, от тех, которые можно интерпретировать как устойчивый стресс или дистресс; это важно для разграничения, например, художественно шокирующего, но функционального воздействия и действительно вредоносного эффекта. В качестве возможной перспективы может быть рассмотрено включение лонгитюдного компонента для оценки длительности сохранения физиологических изменений и их связи с самоотчётами о состоянии. Описанное выше носит проектный характер и подлежит апробации в ходе экспериментов.

Практическая значимость предполагаемого исследования заключается в создании научно обоснованной концепции и потенциального прототипа системы оценки нейрофизиологического воздействия цифрового контента, которую в дальнейшем можно использовать как вспомогательный инструмент в рамках психофизиологической или психологической судебной экспертизы. Важно подчеркнуть, что результаты такого рода способны доба-

вить в обсуждение измеряемое звено: **вместо абстрактного тезиса о том, что «контент якобы наносит вред», можно опираться на данные о том, что в типичной группе пользователей данный класс стимулов вызывает устойчивую, статистически значимую гиперактивацию маркеров стресса, сопровождающуюся, к примеру, нарушением когнитивного контроля по данным ЭЭГ и субъективными жалобами**. С учетом накопленного массива работ, демонстрирующих как возможности, так и ограничения подхода к измерению эмоциональных реакций на медиаконтент, исследование представляется научно обоснованным, методологически реализуемым и способным заложить фундамент для дальнейшей интеграции нейрофизиологических данных в трансдисциплинарные модели оценки риска от цифрового контента.

Принципиальное значение имеет разграничение научной, экспертной и правовой плоскостей анализа. Даже если определенный класс аудиовизуальных стимулов вызывает у части испытуемых устойчивую стрессовую активацию или ухудшение показателей когнитивного контроля, это само по себе не означает, что такой контент должен быть признан противоправным. Квалификация требует самостоятельного анализа правовой нормы, охраняемого интереса, характера общественной опасности, адресата воздействия, контекста распространения материала, возрастной категории аудитории и соразмерности возможного ограничения прав. Следовательно, нейрофизиологические данные могут использоваться только как вспомогательный фактический материал в рамках комплексной экспертизы, но не как самостоятельное основание для запрета, блокировки или иной меры ограничения оборота контента.

Отдельного рассмотрения требуют теоретико-правовые и этические основания возможного использования нейрофизиологических данных в контексте регулирования распространения контента. Отношение различных юрисдикций к нейроданным остается крайне осторожным: подчеркиваются как методологические ограничения фМРТ и смежных технологий, так и опасность их гиперболизации в глазах правоприменителей. В обзорах по рассматриваемой проблематике указывается, что существующие инструменты характеризуются ограниченной пространственно-временной разрешающей способностью, значительной вариативностью (проще говоря, у каждого человека по-своему) и высокой чувствительностью к контексту, что влечет риски некорректной интерпретации и завышенных ожиданий при их использовании в судебных разбирательствах.

Близкий круг вопросов обсуждается также и в области этики и исследований, посвященных «нейроправам», защите когнитивного суверенитета. Авторы, анализирующие последствия распространения нейротехнологий, подчеркивают, что доступ к данным о мозговой активности и их использование для целей оценки поведения, предпочтений или убеждений человека затрагивают базовые права на неприкосновенность частной жизни, свободу мысли и свободу от дискриминации. Доклады международных организаций и научные исследования подчеркивают необходимость разработки регулятивных рамок, учитывающих специфику нейротехнологий, их двойное назначение и возможное коммерческое и государственное злоупотребление.

Ключевым ограничением предлагаемой методики выступает неизбежная многозначность связи между физиологическими маркерами и психологическими интерпретациями. На уровне психофизиологии давно обсуждается, что одни и те же показатели активации – например, повышение электродермальной активности или изменение сердечного ритма – могут отражать как негативно окрашенный стресс, так и познавательное заинтересованное возбуждение, зависящее от мотивации и контекста. Аналогично, фронтальная ЭЭГ-асимметрия, традиционно трактуемая в терминах «приближение-избегание», проявляет сложные зависимости от типа задачи, социального контекста и индивидуальных особенностей. Обзоры психофизиологических исследований медиа подчеркивают, что без строгого экспериментального контроля и сопоставления с поведенческими и самоотчетными данными риск переинтерпретации отдельных физиологических индикаторов остаётся высоким. Поэтому в рамках исследования предполагается изначально закладывать многокомпонентную модель оценки, в которой физиологические показатели рассматриваются совместно с поведенческими метриками (качество выполнения когнитивных задач, изменение внимания и т.п.), что позволяет существенно снизить двусмысленность интерпретаций.

Отдельной задачей, имеющей как научное, так и практическое значение, является проблема повышенной убедительности объяснений, содержащих ссылки на нейроданные и изображения мозга, даже если их добавление не повышает реальной объяснительной силы аргументации. Экспериментальные работы показывают, что визуализация мозговой активности (например, в виде фМРТ-снимков) способна усиливать доверие неспециалистов к представленным доказательствам и объяснениям, создавая видимость повышенной объективности. В связи с этим любые разработки в области применения нейрофизиологических методов к оценке контента должны сопровождаться четким описанием границ валидности и статистической неопределённости получаемых результатов, а также коммуникационной стратегией, минимизирующей риск их некритичного восприятия регуляторами, судьями и широкой публикой. Предполагается снижать этот риск за счет прозрачного изложения методики, обязательного указания доверительных интервалов и демонстрации отрицательных результатов и тщательной валидации моделей на независимых выборках.

Таким образом, предлагаемое исследование занимает промежуточное положение между фундаментальными трудами ученых, изучающих медиа, и прикладными задачами правового регулирования. С одной стороны, он опирается на развитую традицию психофизиологических исследований медиа, демонстрирующих, что динамика сердечно-сосудистых показателей, EDA и ЭЭГ надёжно отражает процессы внимания, эмоционального реагирования и когнитивной нагрузки при взаимодействии с цифровым контентом. С другой стороны, он учитывает современные дискуссии в нейроправе и этике, подчеркивающие как потенциальную полезность нейроданных в юридическом контексте, так и необходимость строгих эпистемологических и правовых ограничений на их использование. Разработанный в рамках возможного исследования протокол оценки нейрофизиологического воздействия контента может быть представлен не как инструмент категорического вынесения вердиктов о «деструктивности», а как научно обоснованный источник дополнительной информации о риске того или иного класса стимулов, дополняющий традиционные критерии. Исследование создает основу для реализации правовых механизмов противодействия деструктивному контенту и установления того, каким образом результаты нейропсихофизиологических измерений могут быть интегрированы в процедуры экспертной оценки контента при сохранении приоритета прав и свобод человека.

В качестве следующего шага представляется целесообразным конкретизировать программу исследования, которая позволил бы не только продемонстрировать принципиальную реализуемость подхода, но и обеспечить возможность обсуждения нормативных предложений. Опираясь на существующие психофизиологие протоколы, где

систематически используются сердечно-сосудистые показатели, электродермальная активность, ЭЭГ и иногда методы нейровизуализации для анализа реакции на аудио- и видеостимулы, можно предложить многофакторный подход с изменениями как типов контента, так и контекстом его предъявления. В идеале набор стимулов должен включать несколько уровней эмоционального воздействия, калиброванных на основе ранее описанных баз изображений и клипов, дополненных авторскими «подозрительными» стимульными наборами, отражающими специфические формы психоделического или трудноинтерпретируемого контента.

Испытуемые в таком исследовании могут быть разделены на несколько групп в зависимости от возраста, уровня медиапотребления и наличия или отсутствия повышенной уязвимости (например, по показателям тревожности или предыдущего негативного опыта взаимодействия с подобным контентом). Подобная стратификация согласуется с данными о том, что индивидуальные различия существенно модифицируют психофизиологические реакции на стимулы и, следовательно, должны учитываться уже на стадии планирования выборки. Эксперимент может включать последовательные блоки просмотра нейтральных, эмоционально стандартизированных (исследователь знает эмоциональный заряд контента) и подозрительных стимулов, разделённых периодами восстановления, в ходе которых регистрируются как физиологические показатели покоя, так и субъективные самоотчёты. Наряду с регистрацией непрерывных сигналов (ЭЭГ, ЭКГ, EDA, дыхание) целесообразно включить поведенческие задания на внимание и когнитивный контроль до и после эксперимента, что позволит оценить возможные краткосрочные изменения когнитивного функционирования под влиянием предъявленного материала.

В рамках эксперимента обработка данных должна будет опираться на методы мультимодальной интеграции. Современные обзоры по мультимодальному распознаванию эмоций подчеркивают, что комбинирование нескольких физиологических каналов (EDA, ЭКГ, ЭЭГ, дыхание и др.), а также, при наличии, поведенческих и визуальных данных, существенно повышает точность классификации эмоциональных состояний по сравнению с моноканальными подходами.

Концептуально интерпретация полученных нейрофизиологических паттернов должна опираться на мотивационно-эмоциональные модели, разработанные в рамках психофизиологии эмоций. Исследования демонстрируют, что эмоциональная обработка стимулов может быть описана в терминах двух основных измерений – валентности и активации, связанных с активацией оборонительных и аппетитивных мотивационных систем, что находит отражение как в мозговых потенциалах, так и в автономных реакциях. Для целей предлагаемого исследования это означает, что «негативность» контента следует операционализировать как систематическое смещение активации в сторону оборонительного контура реакции при высокой интенсивности возбуждения. Дополнительно возможен анализ эффектов повторного показа: часть работ указывает на изменение как субъективных оценок, так и физиологических реакций при многократном предъявлении эмоциональных стимулов, включая снижение аффективного ответа и модификацию событий-связанных потенциалов, что важно для обсуждения возможных эффектов десенсибилизации или, напротив, сенсибилизации к определённым типам контента. Предлагаемая интерпретационная рамка является предположением и должна быть проверена в ходе будущего эксперимента.

Если гипотеза будет подтверждена, ожидаемые результаты могут быть описаны как выявление диапазона профилей нейрофизиологического реагирования, характерных для разных классов контента. Можно выдвинуть предположение, что для части психоделических или трудноинтерпретируемых стимулов будут обнаружены паттерны, включающие сочетание высокой автономной активации (по EDA и сердечному ритму), усиления индексов когнитивной нагрузки и возможного ухудшения показателей когнитивного контроля в поведенческих задачах. Однако столь же вероятно, что часть подобных стимулов окажется ближе к профилю поискового или познавательного возбуждения, описанному в традиции исследований ориентировочного рефлекса и естественного селективного внимания, что ещё раз подчеркнёт невозможность прямого отождествления физиологической активации с «вредом» без учёта контекста и субъективной интерпретации. Этот выстраиваемый спектр реакций может быть представлен в виде индексов риска, в рамках которых конкретные классы контента занимают различные области, а не делятся на бинарные категории «допустимый/запрещенный».

Понимание ограничений и возможных направлений развития предлагаемого подхода тесно связано с общими тенденциями в области анализа мультимодальных эмоциональных данных. Современные обзоры показывают, что, несмотря на впечатляющий прогресс в точности распознавания эмоциональных состояний, сохраняются серьезные проблемы с обобщением моделей на новые выборки, условия и культурные контексты. В исследовании отмечается, что без систематической работы по сбору разнообразных данных и открытой публикации кодов и моделей существует риск переоценки реальной надёжности систем. Для нейрофизиологических оценок воздействия контента это означает необходимость многоцентровых исследований, репликаций и, по возможности, создания открытых баз анонимизированных сигналов, полученных при просмотре стандартизированных наборов стимулов. Только в этом случае можно говорить о формировании действительно устойчивых и проверяемых критериев, которые могли бы претендовать на использование в экспертной практике.

Важно подчеркнуть, что предлагаемая интеграция нейропсихофизиологических методов в оценку контента должна рассматриваться не как подмена существующих подходов, а как их дополнение. Нейрофизиологические данные способны внести в анализ измеримое, количественно описываемое звено, показывая, каким образом тот или иной класс стимулов влияет на эмоциональное состояние, внимание и когнитивные процессы типичного пользователя. При этом с нормативной точки зрения фундаментальными остаются выводы о недопустимости использования нейротехнологий для скрытой манипуляции, дискриминации или принуждения. В этой связи даже успешно валидированный протокол оценки нейрофизиологического воздействия контента может быть оправдан только в формате строго регламентированной, прозрачной и подотчетной процедуры, ориентированной на защиту пользователей и соблюдение их прав, а не на расширение возможностей контроля над ними.

Таким образом, в настоящей статье предложено теоретическое обоснование гипотезы о возможности использования нейрофизиологических показателей для оценки влияния аудиовизуального контента. На основе анализа литературы сформулированы ключевые методологические принципы, которые должны лечь в основу будущего эмпирического исследования. Современные исследования в области психофизиологии, аффективных вычислений и нейромаркетинга демонстрируют принципиальную возможность регистрации и количественного описания эмоциональных, стрессовых и когнитивных реакций на

аудиовизуальные стимулы с использованием ЭЭГ, показателей сердечно-сосудистой системы, электродермальной активности и других каналов. На этом фоне предложенная в статье концепция интегральной оценки нейрофизиологического профиля контента, в том числе невербального и психоделического, представляет собой логическое продолжение уже существующих методологических и технологических разработок, адаптированное к задаче экспертизы трудноподдающегося лингвистическому анализу материала.

Сформулированная исследовательская программа предполагает переход от качественных рассуждений о «вреде» или «деструктивности» контента к операционализируемым показателям уровня активации оборонительных мотивационных систем, выраженности негативной валентности, когнитивной нагрузки и возможных краткосрочных изменений в работе систем внимания и контроля. При этом принципиально важно, что речь идет не о введении жесткой бинарной классификации «допустимый/запрещенный» на основе одного-двух физиологических индексов, а о построении шкалы нейрофизиологического риска, на которой различные классы стимулов занимают свои области. Такой подход, с одной стороны, обеспечивает большую гибкость и чувствительность к нюансам контента и контекста его предъявления, а с другой – позволяет встроить нейрофизиологические данные в уже существующие системы экспертной оценки, не подменяя, а дополняя психологический, лингвистический и юридический анализ.

Одновременно проведённый в статье анализ теоретико-методологических и правовых ограничений ясно показывает, что использование нейрофизиологических методов в контексте регулирования контента не может рассматриваться как нейтральный технический инструмент. Многозначность связи между физиологическими маркерами и психологическими состояниями, межиндивидуальная и межкультурная вариативность реакций задают жёсткие рамки для любых практических предложений.

Перспективы дальнейших исследований в обозначенном направлении связаны, во-первых, с проведением многоцентровых эмпирических исследований с участием разнообразных по возрасту, культурному фону и уровню медиапотребления выборок людей, что необходимо для проверки обобщаемости предлагаемой модели. Во-вторых, представляется важным развивать методы, повышающие устойчивость классификаторов к вариативности условий регистрации и составов выборок, а также разрабатывать открытые базы анонимизированных нейрофизиологических данных, полученных при просмотре стандартизированных наборов стимулов. В-третьих, требуется дальнейшая междисциплинарная работа на стыке нейронауки и права, направленная на выработку четких критериев допустимости и границ ответственности при использовании нейрофизиологических данных. В совокупности эти шаги могут привести к формированию научно обоснованного, методологически прозрачного подхода к оценке воздействия контента, в котором возможности нейронауки используются как дополнительный источник данных для более взвешенных и ответственных решений в интересах общества.

Литература

1. Адиатуллин А.В., Башкович А.Б., Шмонов М.М., Галкина Н.В. Разработка алгоритма распознавания эмоциональных реакций при просмотре рекламных роликов по данным ЭЭГ // Практический маркетинг. – 2020. – № 4 (278). – С. 3 – 10.
2. Вайсертрейгер А.С.-Р. Комплексное исследование динамики физиологических показателей у человека при восприятии сенсорных эмоциогенных стимулов: дис. ... канд. биол. наук. Санкт-Петербург, 2018. 192 с.
3. Горобец Т.Н., Ковалев В.В., Огнивов В.В. Влияние визуальных и аудиальных стимулов на функциональные состояния и работоспособность. // Акмеология. – 2015. – № 3 (55). – С. 194 – 195.
4. Ковальчук М.В., Яцишина Е.Б., Демин В.А., Холодный Ю.И. Необходимость правового регулирования применения нейрокогнитивных и психофизиологических технологий в практике // Вестник МГПУ. Серия: «Юридические науки». – 2022. – Т. 3. – № 47. – С. 85 – 97.
5. Кожевников С.П., Лобанова О.С., Саладовникова Е.Н. Влияние компьютерных игр с элементами насилия на изменение спектров мощности ЭЭГ. // Вестник Удмуртского университета. Серия «Биология. Науки о Земле». – 2012. – № 2. – С. 56 – 61.
6. Полубинская С.В. Использование данных нейронаук в доктрине уголовного права и судебной практике // Труды Института государства и права Российской академии наук. – 2019. – № 14 (5). – С. 9 – 37.
7. Baig M.Z., Kavakli M. (2019). A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. Multimodal Technologies and Interaction. Vol. 3. No. 2. 20 p.
8. Baldo D. et al. (2022). The heart, brain, and body of marketing: Complementary roles of neurophysiological measures in tracking emotions, memory, and ad effectiveness. Psychology & Marketing. No. 39. Pp. 1979-1991.
9. Bartholow B., Bushman B., Sestir M. (2006). Chronic violent video game exposure and desensitization to violence: Behavioral and event-related brain potential data. Journal of Experimental Social Psychology. No. 42. Pp. 532-539.
10. Braithwaite J.J., Watson D.G., Jones R. and Rowe M. (2013). A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRS) for Psychological Experiments. Technical Report, 2nd version: Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre, University of Birmingham, UK. Pp. 1-43.
11. Brown C.M.L. (2024). Neurorights, Mental Privacy, and Mind Reading. Neuroethics. Vol. 17. No. 2. P. 34.
12. Brown T.R., Murphy E.R.D. (2010). Through a scanner darkly: Functional neuroimaging as evidence of a criminal defendant's past mental states. Stanford Law Review. Vol. 62. No. 6. Pp. 1115-1172.
13. Cartocci G. et al. (2017). Electroencephalographic, Heart Rate, and Galvanic Skin Response Assessment for an Advertising Perception Study: Application to Antismoking Public Service Announcements. J Vis Exp, (126), e55872. 9 p.
14. Engelhardt C.R., Bartholow B.D., Kerr G.T., Bushman B.J. (2011). This Is Your Brain on Violent Video Games: Neural Desensitization to Violence Predicts Increased Aggression Following Violent Video Game Exposure. Journal of Experimental Social Psychology. No. 47. Pp. 1033-1036.
15. Feigenson N. (2006). Brain imaging and courtroom evidence: on the admissibility and persuasiveness of fMRI. International Journal of Law in Context. Vol. 2. No. 3. Pp. 233-255.
16. Fujimura T., Katahira K., Okanoya K. (2013). Contextual modulation of physiological and psychological responses triggered byemotional stimuli. Front. Psychol. Vol. 4. Art. 212. 7 p.
17. Goodson S., Turner K.J., Pearson S.L., Carter P. (2021). Violent Video Games and the P300: No Evidence to Support the Neural Desensitization Hypothesis. Cyberpsychology, behavior and social networking. Vol. 24. No. 1. Pp. 48-55.
18. Harper E., McCay A. (2025). Research Brief: Neurotechnology – Integrating Human Rights in Regulation

2025 / The Geneva Academy of International Humanitarian Law and Human Rights. 32 p.

19. Ienca M., Andorno R. (2017). Towards new human rights in the age of neuroscience and neurotechnology. Life Sci Soc Policy. Vol. 13. No. 5. Pp. 5-27.

20. Lang P.J., Bradley M.M. (2010). Emotion and the motivational brain. Biological psychology. Vol. 84. No. 3. Pp. 437-450.

21. Mathews V.P. et al (2005). Media violence exposure and frontal lobe activation measured by functional magnetic resonance imaging in aggressive and nonaggressive adolescents. Journal of computer assisted tomography. Vol. 29. No. 3. Pp. 287-292.

22. Mercado-Diaz L.R., Veeranki Y.R., Large E.W., Posada-Quintero H.F. (2024). Fractal Analysis of Electrodermal Activity for Emotion Recognition: A Novel Approach Using Detrended Fluctuation Analysis and Wavelet Entropy. Sensors, 24. 17 p.

23. Pan Y. (2025). Emotion Recognition Algorithm Based on Multi-Modal Physiological Signal Feature Fusion Using Artificial Intelligence and Deep Learning // International Journal of Advanced Computer Science and Applications (IJACSA). Vol. 16. No. 6. Pp. 105-118.

24. Perry Fordson H., Xing X., Guo K., Xu X. (2022). Emotion Recognition With Knowledge Graph Based on Electrodermal Activity. Front. Neurosci. Vol. 16. Art. 911767. 12 p.

25. Ravaja N. (2004). Contributions of Psychophysiology to Media Research: Review and Recommendations. Media Psychology. Vol. 6. No. 2. Pp. 193-235.

26. Roshdy A., Karar A., Kork S.A., Beyrouthy T., Nait-ali, A. (2024). Advancements in EEG Emotion Recognition: Leveraging Multi-Modal Database Integration. Appl. Sci, 14. 21 p.

27. Soleymani M., Lichtenauer J., Pun T., Pantic M. (2012). «A Multimodal Database for Affect Recognition and Implicit Tagging». IEEE Transactions on Affective Computing. Vol. 3. No. 1. Pp. 42-55.

28. Wu Y., Mi Q., Gao T.A. (2025). Comprehensive Review of Multimodal Emotion Recognition: Techniques, Challenges, and Future Directions. Biomimetics, 10. 24 p.

29. Recommendation of the Council on Responsible Innovation in Neurotechnology, OECD/LEGAL/0457 (11/12/2019) // URL: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0457 (дата обращения: 13.11.2025).

30. Report of the International Bioethics Committee of UNESCO (IBC) on the ethical issues of neurotechnology // URL: https://unesdoc.unesco.org/ark:/48223/pf0000378724#:~:text=ETHICAL%20ISSUES%20OF%20NEUROTECHNOLOGY%20I,devices%20and%20procedures%20collectively%20named (дата обращения: 13.11.2025).

The potential of neurotechnology in assessing potentially illegal audiovisual content: a hypothesis justification

Rashitkhanov R.S.

Kutafin Moscow State Law University (MSAL)

The article substantiates the scientific validity of the hypothesis about the possibility of using neurophysiological indicators to objectively assess the negative impact of audiovisual content on a person's psychophysiological state and, consequently, to take legal measures to restrict the circulation of such content. The article examines the theoretical, legal, and ethical limitations of using neurophysiological data in the context of regulating digital content. It also substantiates the auxiliary role of neurotechnology as a source of additional information about the risk of content, which complements traditional psychological, linguistic, and legal expert approaches but does not replace them. The concept of empirical verification of a scientific hypothesis is described: the study creates a basis for implementing legal mechanisms to counter destructive content and establishing how the results of neuropsychophysiological measurements can be integrated into expert content assessment procedures while maintaining

Keywords: neurotechnology, neurophysiological indicators, EEG, audiovisual content, illegal content, neurolaw, cognitive sovereignty.

References

1. Adiatullin AV, Bashkovich AB, Shmonov MM, Galkina NV. Development of an algorithm for recognizing emotional reactions when watching commercials based on EEG data // Practical Marketing. - 2020. - No. 4 (278). - Pp. 3 - 10.
2. Weissertreyger AS-R. Comprehensive study of the dynamics of physiological indicators in humans during the perception of sensory emotiogenic stimuli: Diss. ... Cand. Biological Sciences. St. Petersburg, 2018. 192 p.
3. Gorobets TN, Kovalev VV, Ognivov VV. The influence of visual and auditory stimuli on functional states and performance. // Acmeology. - 2015. - No. 3 (55). – P. 194 – 195.
4. Kovalchuk M.V., Yatsishina E.B., Demin V.A., Kholodny Yu.I. The Need for Legal Regulation of the Application of Neurocognitive and Psychophysiological Technologies in Practice // Bulletin of Moscow State Pedagogical Univ. Series: "Legal Sciences". – 2022. – Vol. 3. – No. 47. – P. 85 – 97.
5. Kozhevnikov S.P., Lobanova O.S., Saladovnikova E.N. The Impact of Violent Computer Games on Changes in EEG Power Spectra // Bulletin of Udmurt University. Series "Biology. Earth Sciences". – 2012. – No. 2. – P. 56 – 61.
6. Polubinskaya S.V. Use of Neuroscience Data in the Doctrine of Criminal Law and Judicial Practice // Transactions of the Institute of State and Law of the Russian Academy of Sciences. – 2019. – No. 14 (5). – P. 9 – 37.
7. Baig M.Z., Kavakli M. (2019). A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. Multimodal Technologies and Interaction. Vol. 3.No. 2.20 p.m.
8. Baldo D. et al. (2022). The heart, brain, and body of marketing: Complementary roles of neurophysiological measures in tracking emotions, memory, and ad effectiveness. Psychology & Marketing. No. 39. pp. 1979-1991.
9. Bartholow B., Bushman B., Sestir M. (2006). Chronic violent video game exposure and desensitization to violence: Behavioral and event-related brain potential data. Journal of Experimental Social Psychology. No. 42. pp. 532-539.
10. Braithwaite J. J., Watson D. G., Jones R. and Rowe M. (2013). A Guide for Analyzing Electrodermal Activity (EDA) & Skin Conductance Re-sponses (SCRS) for Psychological Experiments. Technical Report, 2nd version: Selective Attention & Awareness Laboratory (SAAL) Behavioral Brain Sciences Centre, University of Birmingham, UK. pp. 1-43.
11. Brown C.M.L. (2024). Neurorights, Mental Privacy, and Mind Reading. Neuroethics. Vol. 17.No. 2. P. 34.
12. Brown T.R., Murphy E.R.D. (2010). Through a scanner darkly: Functional neuroimaging as evidence of a criminal defendant's past mental states. Stanford Law Review. Vol. 62.No. 6.Pp. 1115-1172.
13. Cartocci G. et al. (2017). Electroencephalographic, Heart Rate, and Galvanic Skin Response Assessment for an Advertising Perception Study: Application to Antismoking Public Service Announcements. J Vis Exp, (126), e55872. 9 p.m.
14. Engelhardt C.R., Bartholow B.D., Kerr G.T., Bushman B.J. (2011). This Is Your Brain on Violent Video Games: Neural Desensitization to Violence Predicts Increased Aggression Following Violent Video Game Exposure. Journal of Experimental Social Psychology. No. 47. pp. 1033-1036.
15. Feigenson N. (2006). Brain imaging and courtroom evidence: on the admissibility and persuasiveness of fMRI. International Journal of Law in Context. Vol. 2.No. 3.Pp. 233-255.
16. Fujimura T., Katahira K., Okanoya K. (2013). Contextual modulation of physiological and psychological responses triggered byemotional stimuli. Front. Psychol. Vol. 4. Art. 212.7 p.
17. Goodson S., Turner K. J., Pearson S. L., Carter P. (2021). Violent Video Games and the P300: No Evidence to Support the Neural Desensitization Hypothesis. Cyberpsychology, behavior and social networking. Vol. 24.No. 1.Pp. 48-55.
18. Harper E., McCay A. (2025). Research Brief: Neurotechnology – Integrating Human Rights in Regulation 2025 / The Geneva Academy of International Humanitarian Law and Human Rights. 32 p.m.
19. Ienca M., Andorno R. (2017). Towards new human rights in the age of neuroscience and neurotechnology. Life Sci Soc Policy. Vol. 13.No. 5.Pp. 5-27.
20. Lang P.J., Bradley M.M. (2010). Emotion and the motivational brain. Biological psychology. Vol. 84.No. 3.Pp. 437-450.
21. Mathews V.P. et al (2005). Media exposure and frontal lobe activation measured by functional magnetic resonance imaging in aggressive and nonaggressive adolescents. Journal of computer assisted tomography. Vol. 29.No. 3.Pp. 287-292.
22. Mercado-Diaz L.R., Veeranki Y.R., Large E.W., Posada-Quintero H.F. (2024). Fractal Analysis of Electrodermal Activity for Emotion Recognition: A Novel Approach Using Detrended Fluctuation Analysis and Wavelet Entropy. Sensors, 24. 17 p.
23. Pan Y. (2025). Emotion Recognition Algorithm Based on Multi-Modal Physiological Signal Feature Fusion Using Artificial Intelligence and Deep Learning // International Journal of Advanced Computer Science and Applications (IJACSA). Vol. 16.No. 6.Pp. 105-118.
24. Perry Fordson H., Xing X., Guo K., Xu X. (2022). Emotion Recognition With Knowledge Graph Based on Electrodermal Activity. Front. Neurosci. Vol. 16. Art. 911767. 12 p.
25. Ravaja N. (2004). Contributions of Psychophysiology to Media Re-search: Review and Recommendations. Media Psychology. Vol. 6.No. 2.Pp. 193-235.
26. Roshdy A., Karar A., Kork S.A., Beyrouthy T., Nait-ali, A. (2024). Advancements in EEG Emotion Recognition: Leveraging Multi-Modal Database Integration. Appl. Sci, 14.21 p.