

# КРИЗИС ВЕРИФИКАЦИИ В ЭПОХУ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ: ОТ УНИВЕРСАЛЬНОЙ ОБЪЕКТИВНОСТИ К ПЕРСОНАЛИЗИРОВАННОЙ ИНТЕРПРЕТАЦИИ

**Неткачев Кирилл Иванович**

Аспирант, Тюменский государственный университет  
belkir0712@gmail.com

## THE VERIFICATION CRISIS IN THE AGE OF RECOMMENDER SYSTEMS: FROM UNIVERSAL OBJECTIVITY TO PERSONALIZED INTERPRETATION

**K. Netkachev**

*Summary:* The present paper offers a conceptual inquiry into the profound deformation of information verification mechanisms catalyzed by contemporary recommendation architectures. The study delineates a transition from the classical paradigm of universal objectivity toward a fragmented model of personalized interpretation, stringently dictated by algorithmic filtering. Central to this investigation are the cognitive and technological drivers behind the crystallization of echo chambers and filter bubbles. Special emphasis is placed on an interdisciplinary analysis of the socio-philosophical repercussions stemming from the algorithmic radicalization of attention and the consequent disintegration of a unified informational space. The concluding section articulates a framework for normative interventions, specifically highlighting the architectural engineering of information diversity and the implementation of Explainable Artificial Intelligence as essential instruments for «demystifying» predictive logic and revitalizing epistemic trust within the digital ecosystem.

*Keywords:* post-truth, verification, recommendation systems, attention economy, filter bubbles, echo chambers, digital regulation.

*Аннотация:* В статье предпринимается попытка концептуального осмысления того, как современные алгоритмы рекомендации деформируют устоявшиеся механизмы верификации данных. Автор фокусирует внимание на переходе от классического идеала универсальной объективности к сугубо фрагментарной модели персонализированной интерпретации, жестко детерминированной работой алгоритмических фильтров. В фокусе исследования находятся когнитивные и технологические драйверы формирования эхо-камер и пузырей фильтров. Особый акцент сделан на междисциплинарном анализе последствий алгоритмической радикализации внимания и сопутствующего распада единого информационного пространства на изолированные сегменты. Финальная часть статьи посвящена обоснованию нормативных контрмер, среди которых выделяются архитектурное проектирование информационного разнообразия и стандартизация объяснимого искусственного интеллекта: данные решения рассматриваются как необходимый инструментарий для «демистификации» предиктивных алгоритмов и реанимации эпистемического доверия в цифровой среде.

*Ключевые слова:* постправда, верификация, рекомендательные системы, экономика внимания, пузырь фильтров, эхо-камеры, цифровое регулирование.

### Введение

#### Парадигмальный сдвиг в потреблении информации и кризис объективности

В современной цифровой среде мы наблюдаем не просто количественный рост данных, но фундаментальный слом того, что можно было бы назвать эвристическими стратегиями. Данный процесс требует теоретического и прежде всего философского осмысления. Если традиционная модель информационного поиска, которую принято называть Pull-моделью, предполагала наличие у пользователя активной или даже «волютаристской» позиции, то нынешняя архитектура глобальных платформ (включая экосистему Meta, чья деятельность признана экстремистской и запрещена в РФ, а также TikTok и сервисы Google) навязывает принципиально иную схему, которую можно обозначить как Push-модель. Информация перестает быть объектом поиска; она превен-

тивно «атакует» субъекта, опираясь на массив его цифровых следов и результаты предиктивного анализа поведения [1, с. 17–18].

Этот тектонический сдвиг спровоцировал глубокую эрозию верификационных практик. Ранее, в рамках классического научного и медийного дискурса, статус «истинного» присваивался определенному утверждению после его прохождения через систему фильтров, контролируемую профессиональными посредниками — так называемыми «привратниками» (gate keepers), чья легитимность подтверждалась независимыми институтами [2, р. 383–384]. Современный кризис доверия к фактам приводит к тому, что информация признается достоверной не из-за эмпирической точности, а в силу ее комплементарности относительно идентичности конкретного потребителя — критерий достоверности фатально дрейфует в сторону эмоционального резонанса. Как отмечает исследовательница современной

политической культуры и медиа, лауреат Пулитцеровской премии, публицист и литературный критик М. Какутани, в актуальной цифровой среде «каждый имеет право на особое мнение, но не на особые факты» — поляризация зашла настолько далеко, что пользователи соцсетей с трудом приходят к соглашению по поводу одних и тех же фактов, получая «подкрепление собственных заведомых мнений» в своих «камерах без окон» [3, с. 17–18]

«Объективность» в логике современных платформ становится экономическим балластом. Технологические гиганты ориентированы на «коэффициент удержания» (*retention rate*), который находится в прямой зависимости от когнитивного комфорта аудитории. Процедура верификации, требующая от пользователя интеллектуальных усилий и зачастую провоцирующая когнитивный диссонанс [4, с. 35–39], враждебна идеалу интерфейсной бесшовности (*seamless interface*) — то есть пользовательского опыта (*user experience design, UX*), при котором взаимодействие с системой происходит максимально гладко, без технических пауз или логических препятствий. Верификация данных — это «плохой UX» с точки зрения удержания внимания, так как она создает «трение» (*friction*). Идеал современного интерфейса — сделать так, чтобы пользователь скользил по контенту, не «спотыкаясь о сомнения». Платформы сознательно минимизируют любое когнитивное «трение», предлагая пользователю контент, заранее откалиброванный под его психографический профиль (то есть внутренние мотивации, убеждения и иные психологические установки субъекта, определяющие его поведение). Объективность в этой системе координат изживается не вследствие цензурных запретов, а в силу рыночной логики: поддержание аффективной вовлеченности монетизируетсякратно эффективнее, чем трансляция доказательного контента. [5, р. 103–105].

Конечным пунктом этого процесса становится дезинтеграция верифицируемого знания как такового — некогда общепринятой совокупности эмпирически проверяемых данных, формирующих информационное пространство или научный дискурс. В качестве философского фундамента этого кризиса уместно рассмотреть тезис Б. Латура о том, что инструменты критики подвергают сомнению сами положения дел (*matters of fact*). Латур отмечает, что «безотлагательный ревизионизм» имитирует процедуры социологии науки, чтобы «разрушить доказательства», выработанные исследовательской практикой. Когда любой эмпирический довод с легкостью клеймится как результат субъективной интерпретации, сама идея верификации теряет смысл [6, с. 31–34]. Это не просто затрудняет коммуникацию, но фактически уничтожает пространство для рациональной дискуссии и какого-либо социального консенсуса [7, р. 158].

## 1. Теоретический базис Эволюция верификации и цифровая реконфигурация доверия

Классическая верификационная модель, доминировавшая в пространстве массовых коммуникаций прошлого века, опиралась на концепцию «привратников» (*gate keepers*). Сам термин, предложенный изначально К. Левином еще в 1943 году [8, р. 40], в дальнейшем (в 1950-е) был адаптирован к медиатеории Д. Уайтом [9, р. 54–56]. Согласно этой логике, редакционные институты (в классической медиатеории середины XX века — это традиционные средства массовой информации (*legacy media*), обладающие выстроенной структурой профессионального отбора и проверки контента) брали на себя роль централизованного фильтра, выступая гарантом эпистемической безопасности публичной сферы. В рамках подобной иерархии истина обладала выраженным институциональным характером. Однако масштабная цифровая трансформация вкупе с переходом к алгоритмическому ранжированию фактически лишила традиционные институты их монополии на подтверждение реальности.

С нашей точки зрения, наиболее релевантной исследовательской оптикой для анализа этой трансформации представляется подход С.-Т. Нгуена, который смещает фокус внимания с содержания информации на структуру каналов связи — с верификации конкретного контента на саму архитектуру доверия [7, р. 142–152]. Нгуен настаивает на принципиальном различии двух конфигураций искажения информационной среды: эпистемических пузырей (*epistemic bubbles*), возникающих вследствие исключения релевантной информации из поля зрения пользователя, и эхо-камер (*echo chambers*), механизмы которых направлены на подрыв доверия к любым внешним источникам. Если эпистемический пузырь является во многом побочным продуктом алгоритмической селекции — «пассивным» фильтром, — то природа эхо-камеры куда более сложна. Этот феномен базируется на системной манипуляции кредитом доверия: здесь альтернативные точки зрения не просто отсутствуют, но и превентивно дискредитируются. Как подчеркивает И. Калпокас, подобные структуры позволяют адептам замыкаться в самодостаточных эпистемических сообществах, где любой внешний довод воспринимается не как повод для рациональной дискуссии, а как подтверждение «повстанческого» статуса их собственного знания перед лицом «враждебного» мейнстрима [10, р. 102].

В этой точке верификация подменяется актом социальной лояльности. Мы наблюдаем симптоматичный переход от фактологической проверки к процедуре группового подтверждения. Как отмечает И. Калпокас, в такой системе координат «ключевым критерием при выборе между конкурирующими заявлениями об исти-

не является то, хотел бы человек (или не хотел бы), чтобы что-то было правдой» [10, р. 18]. Аудитория ожидает удовлетворения своих «сокровенных побуждений независимо от содержания сообщения», а коммуникаторы навязывают им Push-модель, используя эту готовность. В итоге «постправда — это не манипуляция, а сговор» между лидерами мнений и их последователями. Нынешняя цифровая архитектура фактически легитимирует подобные когнитивные стратегии, поскольку они минимизируют дискомфорт субъекта при одновременной максимизации его вовлеченности. Рекомендательные системы и механизмы ранжирования преследуют иные цели: не установление истины, а тотальное удержание внимания. Объективность в такой системе координат изживается как избыточная категория, уступая место приему ожидания над содержанием.

Переориентация субъекта с поиска истины на подтверждение групповой идентичности не является случайным сбоем в системе коммуникаций. Напротив, повторимся, эта когнитивная трансформация находит мощную технологическую поддержку в архитектуре современных медиаплатформ. Если эхо-камеры создают структурные условия для изоляции, то сами механизмы функционирования цифровой среды превращают эту изоляцию в товар. Таким образом, анализ архитектоники доверия неизбежно переходит в плоскость исследования прагматики интерфейсов, где конфликт между вовлеченностью и достоверностью становится системным фактором.

## 2. Механизмы эрозии: вовлеченность против достоверности

Неразрешимый конфликт между требованием эффективности и внутренней логикой рекомендательных систем укоренен в самом фундаменте бизнес-моделей ИТ-гигантов. В условиях господства экономики внимания информация превращается в инструмент аффективного захвата. Поскольку ключевым мерилем эффективности алгоритма выступает не объективная точность, а глубина вовлеченности, фактологическая достоверность данных неизбежно приносится в жертву их способности провоцировать мгновенную реакцию.

Когда индивид фиксирует, что некое утверждение внутри его сегментированного информационного поля аккумулирует критическую массу одобрительных реакций, его психика склонна интерпретировать это как свидетельство достоверности. Платформы лишь цементируют такой эффект, превентивно ограничивая трансляцию диссонирующей информации. Классический психологический механизм социального доказательства (*social proof*), основательно исследованный в рамках экспериментальной социальной психологии уже к середине 1980-х [11, с. 254], получает в цифровой архитектуре мощный рычаг масштабирования.

Группа исследователей под руководством С. Арала эмпирически подтвердила, что недостоверные сообщения (фейки) циркулируют в сетях в среднем в шесть раз быстрее правды [12, р. 1147–1148]. Причина кроется в специфической природе фальсификаций: они обладают более выраженной новизной и аффективным зарядом. Так процедура верификации окончательно мимикрирует под персонализированную интерпретацию. Данные, не инкорпорированные в ценностную решетку конкретной эхо-камеры, подвергаются активной дискредитации.

3. Түфекчи описывает современную стратегию контроля над информацией, называя ее «цензурой через шум» (*ensorship through noise*) и «замутнением онлайн-вод» (*muddy the online waters*) [13, р. 29] — так за дезинформаций и/или переизбытком информации алгоритмы эффективно скрывают неудобные факты.

В итоге цифровая архитектура перестает быть нейтральным проводником реальности. Ситуация достигает апогея на фоне экспансии генеративных моделей и технологий дипфейков, когда даже непосредственное визуальное наблюдение не дает возможность произвести верификацию.

## 3. Алгоритмическая радикализация и фрагментация реальности

Одной из деструктивных характеристик современных рекомендательных механизмов выступает их тенденция к последовательной радикализации пользовательских установок. Процесс этот, известный в научном обиходе как «эффект кроличьей норы» (*rabbit hole effect*), представляет собой технологическую воронку: стремясь к максимизации внимания пользователя, алгоритм старается повышать градус экстремальности контента. В рамках этой логики достоверность сообщения легитимируется способностью резонировать с априорными подозрениями субъекта.

Согласно позиции 3. Түфекчи, алгоритм здесь берет на себя роль агента социального инжиниринга [13, р. xxix]. Зафиксировав пусть даже умеренный интерес пользователя к определенной теме, система начинает агрессивно выстраивать траекторию его потребления в сторону деструктивного скепсиса по отношению к официальным институтам. Современные рекомендательные архитектуры выступают активным множителем социальной поляризации — результатом этого становится распад общества на герметичные фактологические вселенные.

Усугубляется ситуация окончательной девальвацией визуальной достоверности. Когда традиционный субъект философии Нового времени лишается своего последнего эпистемического основания — достоверного чувственного опыта, — процедура верификации окон-

чательно вырождается в акт декларации лояльности референтной группе. Таким образом, дрейф от фильтрации истины к генерации конформной лжи оказывается системным свойством текущей цифровой среды.

#### 4. Философская рефлексия: децентрализация истины и кризис коммуникации

Совокупность описанных выше трансформаций приводит к эпистемологическому кризису: категория истины утрачивает претензию на универсальность, приобретая черты субъективного акта потребления. В терминологии Ж. Бодрийера такое положение дел описывается как триумф гиперреальности: господство алгоритмических симулякров — репрезентаций, не имеющих под собой фактического основания, — в восприятии субъекта оказывается куда более убедительным, нежели сама реальность [14, с. 20–23; 111–112; 147–148].

Подобная архитектура познания провоцирует специфическую форму эпистемического сопротивления. Демонстративный отказ от признания объективных данных зачастую интерпретируется пользователем как акт защиты суверенной интеллектуальной автономии. Истина (или то, что за нее принимается), как уже было сказано, превращается в маркер групповой принадлежности.

Фатальная эрозия верификационных процедур ведет к деградации публичной коммуникации. Если следовать логике Ю. Хабермаса, то непременным условием реализации рационального дискурса выступает наличие разделяемого всеми жизненного мира (*Lebenswelt*) — горизонта фактов, чья истинность не оспаривается [15, с. 539–573]. Однако алгоритмическая фрагментация планомерно уничтожает этот фундамент. Продуктивная дискуссия замещается столкновением аффективных идентичностей. Верификация в цифровую эпоху все чаще мигрирует в плоскость парасоциального взаимодействия: доверие делегируется не абстрактному научному знанию, а конкретным инфлюенсерам. В конечном итоге мы наблюдаем кристаллизацию режима пост-объективности, превращающего публичную сферу в конгломерат несовместимых интерпретаций.

#### Заключение: от реактивной борьбы к архитектурным решениям

Из нашего анализа следует, что в условиях экспансии рекомендательных систем процедура установления достоверности перестала быть волевым актом познающего субъекта. Центральный вызов так называемой «эпохи постправды» состоит в противоречии, заложенном в самой архитектонике цифровых платформ, где рыночный императив удержания внимания вступает в конфликт с сохранением доступа к верифицируемой информации.

Перспективный выход может заключаться не в реактивном преследовании фейков, а в превентивной трансформации информационной среды посредством архитектурного обеспечения серендипности (*serendipity by design*). В данном контексте под серендипностью понимается проектирование таких интерфейсных решений, которые намеренно создают условия для контролируемого столкновения пользователя с релевантным, но незапрашиваемым контентом, выходящим за пределы его сложившихся предпочтений. В отличие от бесшовных рекомендательных моделей, стремящихся к максимальной релевантности, архитектура серендипности восстанавливает «информационное трение», возвращая в цифровую среду элемент продуктивной случайности. На законодательном мировом уровне такой подход уже закреплен в Законе ЕС о цифровых услугах (*Digital Services Act, DSA*), обязывающий крупнейшие онлайн-платформы предлагать пользователям как минимум одну систему рекомендаций, не основанную на профилировании [16, Article 38]. Данная норма фактически легитимизирует право субъекта на получение информации вне рамок его прошлых предпочтений. На практике это реализуется через обязательное внедрение лент, не использующих предиктивный анализ, что позволяет минимизировать эффекты радикализации и фрагментации реальности.

Внедрение в интерфейс элементов контролируемой случайности (например, блоков «альтернативная точка зрения» или лент с низкой степенью персонализации) расширяет когнитивный горизонт пользователя. Это создает условия для восстановления общего коммуникативного пространства, где разные социальные группы вновь начинают оперировать единым набором фактов, пусть и сохраняя право на их различную интерпретацию. Таким образом, серендипность выступает не просто дизайнерским решением, но и инструментом реабилитации объективной реальности в глазах субъекта.

Примером превентивного регулирования также может выступать принцип проектируемой безопасности и прозрачности (*transparency by design*). Речь о внедрении так называемого объяснимого искусственного интеллекта (*Explainable AI*). Суть такого подхода в том, что платформы обязаны встраивать инструменты самоконтроля пользователя в интерфейс, обеспечивая доступность механизмов верификации для устранения непрозрачности алгоритмического отбора. Например, согласно нормам Закона ЕС об искусственном интеллекте (*EU AI Act*), AI-системы должны проектироваться таким образом, чтобы обеспечивать проверяемый уровень прозрачности их функционирования [17, Art. 13(1)]. В контексте социальных медиа это выражается в требовании раскрывать параметры, на основе которых контент был доставлен пользователю. Подобная «демистификация» алгоритмов позволяет восстановить критическую дистанцию, превращая потребление информации в осознанную навигацию.

Практическая польза объяснимого искусственного интеллекта (ХАИ) выражается в праве пользователя на получение содержательного обоснования: например, почему конкретная новость или мнение были помещены в топ выдачи. Когда алгоритм эксплицирует свои критерии (указывая на «высокую популярность в сегментированной группе», «аффективный характер заголовка» и т.п.), он тем самым лишает сообщение его «манипулятивной силы». Это позволяет перейти от пассивного потребления контента к его критическому анализу.

Подводя итог, можно утверждать, что преодоление кризиса доверия в цифровую эпоху невозможно через простой возврат к классической модели «привратников». Современная стратегия защиты информационной среды должна опираться на синтез правового регулирования (DSA, AI Act) и осознанного проектирования

когнитивно-дружелюбных интерфейсов. Вместо «бесшовного» поглощения данных, архитектурные решения будущего должны предлагать модели дизайна с «управляемым сопротивлением», направленные на деавтоматизацию восприятия и восстановление критической дистанции.

В конечном итоге цель подобной трансформации — восстановление эпистемического фундамента публичной сферы и целостности верифицируемой среды. И это представляется возможным через синергию технологического проектирования и правового регулирования. Переход к архитектуре осознанности и прозрачности алгоритмов дает шанс преодолеть «смерть истины», заменив герметичные эхо-камеры открытым, хотя и сложно организованным, пространством верифицируемого знания.

#### ЛИТЕРАТУРА

1. Зубофф Ш. Эпоха надзорного капитализма. Битва за человеческое будущее на новых рубежах власти / пер. с англ. А.Ф. Васильева. — М.: Изд-во Института Гайдара, 2022. — 784 с.
2. White D.M. The «Gate Keeper»: A Case Study in the Selection of News // *Journalism Quarterly*. — 1950. — Vol. 27, № 4. — P. 383–390.
3. Какутани М. Смерть правды / пер. с англ. Л. Сумм. — М.: Эксмо, 2019. — 311 с.
4. Фестингер Л. Теория когнитивного диссонанса / пер. с англ. А.А. Анистратенко, И.В. Знаешевой. — М.: Эксмо, 2018. — 256 с.
5. McIntyre L. Post-Truth. — Cambridge, MA: MIT Press, 2018. — 216 p.
6. Латур Б. Почему критика выдохлась? От фактов к вопросам, вызывающим озабоченность / пер. с англ. П. Маркиной // *Логос*. — 2023. — Т. 33, № 5. — С. 29–61.
7. Nguyen C.T. Echo Chambers and Epistemic Bubbles // *Episteme*. — 2020. — Vol. 17, № 2. — P. 141–161. [Электронный ресурс]. — URL: <https://www.cambridge.org/core/journals/episteme/article/abs/echo-chambers-and-epistemic-bubbles/5D4AC3A808C538E17C50A7C09EC706F0> (дата обращения: 01.02.2026). — DOI: 10.1017/epi.2018.32.
8. Lewin K. Forces Behind Food Habits and Methods of Change // *The Problem of Changing Food Habits: Report of the Committee on Food Habits 1941–1943*. — Washington, DC: National Research Council, National Academy of Sciences, 1943. — Bulletin No. 108. — P. 35–65. (Электронный ресурс). — URL: [nationalacademies.org/read/28634/chapter/4#60](https://www.nationalacademies.org/read/28634/chapter/4#60) (дата обращения: 01.02.2026).
9. Napoli P.M. Social Media and the Public Interest: Media Regulation in the Disinformation Age. — New York: Columbia University Press, 2019. — 296 p.
10. Kalpokas I. A Political Theory of Post-Truth. — Cham: Palgrave Pivot, 2019. — 135 p.
11. Чалдини Р. Психология влияния. 7-е расшир. изд. / пер. с англ. О. Епимахова, И. Селивановой. — М.: Эксмо, 2022. — 663 с.
12. Vosoughi S. The spread of true and false news online / S. Vosoughi, D. Roy, S. Aral. — Текст: электронный // *Science*. — 2018. — Vol. 359, iss. 6380. — P. 1146–1151. — URL: <https://www.science.org/doi/10.1126/science.aap9559> (дата обращения: 01.02.2026). — DOI: 10.1126/science.aap9559.
13. Tufekci Z. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven: Yale University Press, 2017. — 360 p.
14. Бодрийяр Ж. Симулякры и симуляция / пер. с фр. А. Качалова. — М.: Издательский дом ПОСТУМ, 2015. — 240 с.
15. Хабермас Ю. Теория коммуникативной деятельности: в 2 т. / пер. с нем. А. К. Судакова. — М.: Весь Мир, 2022. — 880 с.
16. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services (Digital Services Act) // EUR-Lex: [сайт]. — URL: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj> (дата обращения: 30.01.2026).
17. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) // EUR-Lex: [сайт]. — URL: Regulation - EU - 2024/1689 - EN - EUR-Lex (дата обращения: 30.01.2026).

© Неткачев Кирилл Иванович (belkir0712@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»