АЛГОРИТМИЧЕН МАРКЕТИНГ, ДИГИТАЛНИ ДВОЙНИЦИ И МАНИПУЛАЦИЯ НА ПОТРЕБИТЕЛСКОТО ПОВЕДЕНИЕ: ЕТИКА И РИСКОВЕ В ДИГИТАЛНАТА ИКОНОМИКА

Algorithmic Marketing, Digital Twins, and the Manipulation of Consumer Behavior: Ethics and Risks in the Digital Economy

Василена Василева[1]
e-mail: vasilena.vasileva@unwe.bg [1]

Резюме

Развитието на алгоритмичните социални платформи и изкуствения интелект доведе до революция в маркетинга, като въведе понятието за дигитален двойник на потребителя — модел, който предсказва предпочитания, решения и емоционални реакции. Тази нова парадигма разширява възможностите на предиктивния маркетинг, но поражда и етични дилеми, свързани с автономията, прозрачността и доверието в дигиталната икономика. Настоящият доклад изследва теоретичните и практически аспекти на алгоритмичния маркетинг, разглеждайки етичните рискове от манипулация, злоупотреба с данни и дигитални дълбоки фалшификации (deepfakes). През последното десетилетие алгоритмите за машинно обучение се превърнаха в двигател на дигиталната икономика. Маркетинговите стратегии вече не се базират на масови послания, а на динамично моделиране на поведението чрез големи данни (Big Data) и невронни мрежи. Понятието „дигитален двойник“ (Digital Twin) — първоначално използвано в инженерството — днес се прилага и спрямо индивидуалните потребители. Както отбелязва Ye et al. (2025), цифровите модели позволяват „биоинспирирано“ наблюдение на поведението, което може да бъде използвано както за устойчиво развитие, така и за контрол.

Ключови думи: алгоритмични социални платформи, изкуствен интелект, маркетинг, дигитален двойник, предиктивен маркетинг, етични дилеми, автономия, прозрачност, доверие, дигитална икономика, алгоритмичен маркетинг, етични рискове, манипулация, злоупотреба с данни, дигитални дълбоки фалшификации (deepfakes), машинно обучение, големи данни (Big Data), невронни мрежи, динамично моделиране на поведението, биоинспирирано наблюдение, устойчиво развитие, контрол.

Abstract

The rise of algorithmic social platforms and artificial intelligence has sparked a revolution in marketing, introducing the concept of the consumer’s digital twin — a model that predicts preferences, decisions, and emotional responses. This new paradigm expands the potential of predictive marketing but also raises ethical dilemmas related to autonomy, transparency, and trust within the digital economy. The present report examines both the theoretical and practical aspects of algorithmic marketing, addressing ethical risks such as manipulation, data misuse, and digital deepfakes. Over the past decade, machine learning algorithms have become the driving force of the digital economy. Marketing strategies are no longer based on mass messaging but on dynamic behavioral modeling through Big Data and neural networks. The concept of the “digital twin” — originally used in engineering — is now applied to individual consumers. As noted by Ye et al. (2025), digital models

[1] Докторант към катедра „Индустриален бизнес“, УНСС, e-mail: vasilena.vasileva@unwe.bg

enable “bio-inspired” behavioral observation, which can be leveraged for both sustainable development and control.

Теоретични основи на алгоритмичния маркетинг

Алгоритмичният маркетинг представлява **интердисциплинарен синтез**, който преплита принципите на икономическата рационалност, когнитивната психология, невронауката и изкуствения интелект, създавайки ново познавателно поле — *икономика на предсказуемото поведение*. Тази нова парадигма надхвърля традиционната логика на пазарното взаимодействие и се превръща в **система за поведенческо програмиране**, при която субектът не просто избира, а бива *поведенчески проектиран*.

Икономическата рационалност и нейната дигитална трансформация

Класическата икономическа теория (Homo economicus) приема, че човекът действа рационално, максимизирайки полезността си. Алгоритмичният маркетинг подлага това допускане на радикална ревизия — не защото рационалността е изчезнала, а защото е **алгоритмизирана**.

Благодарение на масиви от поведенчески данни (т.нар. *digital exhaust*), машините вече не предполагат рационалност, а **изчисляват вероятността от определено поведение**. Както показват Tucker (2022) и Kim & Zhang (2024), съвременните маркетингови модели използват Бейесови вероятности и невронни мрежи, за да прогнозират микрорешения — като моментите на колебание преди покупка, или склонността към лоялност към бранд.

Така икономическата рационалност се заменя от *статистическа предсказуемост* — рационалното вече не е норма, а функция на изчислен риск.

Поведенческа икономика: от „убеждаване“ към „архитектура на избора“

Thaler & Sunstein (2021) въвеждат концепцията за *nudge* — леко „подбутване“, което насочва поведението, без да ограничава избора. Алгоритмичният маркетинг е еволюцията на този принцип, но в хипертрофирала форма: докато традиционният *nudge* е съзнателен и целенасочен, **алгоритмичният nudge е автоматизиран, невидим и адаптивен**.

Системите използват модели на когнитивни изкривявания — като ефекта на потвърждението, социалното доказателство или страха от пропускане (FOMO) — за да създадат контексти, в които потребителят възприема предварително програмирания избор като собствен. Това вече не е маркетинг, а *архитектура на поведенческия свят*, в който автономията се преживява, но не се упражнява.

Невроикономика и невромаркетинг: декодиране на съзнанието

Невроикономиката разглежда как мозъкът оценява стойността, удоволствието и очакванията при вземане на решения. Plassmann et al. (2020) доказват, че определени мозъчни региони — например вентромедиалният префронтален кортекс — активират предсказуеми модели на реакция при възприемане на брандове. Алгоритмичният маркетинг използва тези данни за създаване на *невроемпатични* системи: ИИ модели, които адаптират рекламното съдържание в реално време според емоционалния тон на гласа, мимиките или сърдечния ритъм, измерен чрез персонални биосензорни устройства.

Това въвежда ново ниво на персонализация, което можем да наречем **афективна персонализация** — маркетинг, който не просто говори „на теб“, а *чрез теб*.

Изкуственият интелект като когнитивна инфраструктура

Изкуственият интелект не е просто инструмент за автоматизация — той е **когнитивен посредник** между маркетинга и потребителя. Чрез дълбоки невронни мрежи (Deep Learning) и генеративни модели (Generative AI) алгоритмите създават „поведенчески хипотези“, които се проверяват в реално време върху милиони индивиди.

Тези системи не се нуждаят от теории за мотивацията — те извличат закономерности от емпирични данни. Както посочва Floridi (2023), това води до „епистемологична трансформация“: знанието за човека вече не е психологическо, а **изчислително**. ИИ системите моделират вероятностите на поведение, превръщайки емоциите в алгоритмични сигнали, а решенията — в изходи на невронни оптимизации.

От комуникация към предсказателна инфраструктура

Традиционно маркетингът се разглежда като процес на **двупосочна комуникация** между бранд и потребител. В епохата на алгоритмичните системи той се превръща в *предиктивна инфраструктура* — мрежа от модели, която не просто отговаря на нуждите, а ги проектира.

Ye et al. (2025) отбелязват, че съвременните модели функционират аналогично на невронни структури — те учат от грешките си, кодират поведенчески шаблони и „синхронизират“ потребителската активност с икономическите цели на платформата.

Следователно целта на маркетинга вече не е да **убеди**, а да **предопредели**: да създаде такава когнитивна и емоционална среда, в която определено поведение изглежда естествено, неизбежно и лично мотивирано.

Критична перспектива

Тази трансформация поставя фундаментален философски въпрос: ако поведението е предсказано и насочено от алгоритми, **остава ли място за свобода на избора**? Алгоритмичният маркетинг, в своята крайна форма, се доближава до концепцията за *алгоритмично управление* (algorithmic governance) — където пазарът, държавата и медиите се сливат в единна система за управление на поведението.

Оттук следва, че изследването на алгоритмичния маркетинг вече не е просто икономическа дисциплина, а форма на *социална философия*, поставяща под въпрос границите на човешката автономия в епохата на изчислимия човек.

Концепцията за дигитален двойник в маркетинга

Понятието *digital twin* възниква в инженерството като модел за симулиране и мониторинг на физически системи – например турбини, самолети или градска инфраструктура. Но с разрастването на дигиталната икономика и социалните платформи, това понятие **прескочи границите на индустриалното инженерство** и навлезе в хуманитарните и социалните науки.

Днес „дигиталният двойник“ вече не е модел на машина, а **модел на човека**, изграден чрез непрекъснато наблюдение, извличане и интерпретация на данни за поведението му в реална и виртуална среда. Тази трансформация бележи исторически обрат – от *механичен реализъм* към *алгоритмичен антропоцентризъм*.

От инженерния модел към когнитивната симулация

Първоначално инженерният дигитален двойник представляваше точна, синхронизирана реплика на физическа система, която реагира на реални данни, за да предвиди бъдещи състояния. Днес маркетинговият дигитален двойник функционира по аналогичен принцип — но неговата „физическа система“ е **индивидът**.

Този двойник не е статично копие, а *динамична и самообновяваща се конструкция*, тъй като се захранва от:

- поведенчески данни (търсения, покупки, навици);
- биометрични сигнали (пулс, глас, микромимики);
- социални взаимодействия (лайкове, коментари, мрежи);
- контекстуални индикатори (време, място, устройство).

Всяко взаимодействие става *сензор*, който обогатява модела и позволява алгоритмична реконструкция на личността.

Епистемология на дигиталния двойник: от наблюдение към предсказание

Ghosh & Lee (2023) изтъкват, че дигиталните двойници не просто описват миналото поведение, а създават **симулативна прогноза** — способност да се изчислява *вероятното бъдещо „аз“*.

Това означава, че маркетинговите алгоритми вече не работят със статистическа извадка от потребители, а с **уникална дигитална личност**, която реагира, учи и се адаптира. Този модел може да:

- предвиди кога потребителят ще изпита нужда (например глад, интерес, тревожност);
- симулира емоционалното му състояние в различни ситуации;
- изчисли оптималния момент за реклама, послание или продуктова оферта.

Така възниква *икономика на вероятностите*, в която стойността на потребителя не се измерва в покупки, а в **предсказуемост**.

Емоционална персонализация и невидима манипулация

В класическия маркетинг персонализацията означаваше съответствие между продукт и интерес. В епохата на дигиталните двойници тя придобива ново измерение — **афективна персонализация**, при която рекламните алгоритми се настройват според моментните емоционални флуктуации на индивида.

Благодарение на AI системи, които разпознават микровиражи на лицето, тембър на гласа или дори ритъма на писане, дигиталният двойник може да „чувства“ в реално време. Рекламата вече не е послание, а **емоционален диалог**, при който системата знае не само *какво* харесваш, а *кога* си най-уязвим към това, което харесваш.

Тук се размива границата между *персонализация* и *манипулация*: когато съдържанието се адаптира към подсъзнателното ти състояние, изборът престава да бъде изцяло твой.

От огледало към посредник: дигиталният двойник като когнитивен агент

Дигиталният двойник вече не е просто „отражение“ на личността, а **активен посредник между реалното и прогнозираното „аз“**. Той не само моделира поведението, но и го **съ-създава** — участва в процеса на вземане на решения, като предлага, подтиква, а понякога и елиминира алтернативи.

Може да се каже, че дигиталният двойник е *когнитивен агент*, който живее между данните и съзнанието — форма на дигитално съ-аз, чието съществуване е алгоритмично детерминирано.

Както отбелязва Helbing (2022), този феномен представлява „нова кибернетика на човека“, при която системите не само се учат от поведението, но и го моделират обратно, формирайки **затворен цикъл на влияние**.

Етични и философски последици

Появата на дигиталните двойници поставя нови въпроси пред етиката и философията на съзнанието:

- **Кой притежава дигиталния двойник?** – личността, компанията, или алгоритъмът, който го е създал?
- **Какво означава идентичност**, когато алгоритмичната версия на теб взема решения по-бързо и по-точно от самия теб?
- **Къде свършва прогнозата и започва манипулацията?**

Zuboff (2019) предупреждава, че дигиталните двойници превръщат човешкия опит в *суровина за предсказание*, като капитализират бъдещето на поведението. Следователно етичният въпрос вече не е само за поверителността на данните, а за **суверенитета на съзнанието**.

Изкуственият интелект като когнитивен посредник в алгоритмичния маркетинг

Изкуственият интелект (ИИ) вече не може да бъде разглеждан единствено като инструмент за автоматизация или оптимизация на маркетингови процеси. Той е **когнитивен посредник** – интелектуална инфраструктура, която интерпретира, предсказва и моделира човешкото поведение чрез изчислителна логика.

Тази роля надхвърля традиционното разбиране за ИИ като „инструмент“ и го поставя в позицията на **партньор в когнитивното производство на реалността**, т.е. в процеса, чрез който компаниите не просто наблюдават, а *създават* социални и поведенчески феномени.

От автоматизация към когнитивна медиация

Класическата автоматизация заместваше човешки труд с машини. Алгоритмичната когнитивна система, обаче, **замества човешко разбиране с изчислителна интерпретация**. Това означава, че ИИ не просто изпълнява маркетингови функции (например таргетиране), а *мисли чрез данни* – разпознава модели, интерпретира контекст и създава хипотези за бъдещи действия на потребителя.

Както отбелязва Floridi (2023), ИИ извършва „епистемологична трансформация“ – преминаване от хуманно към изчислително познание. Това е преход от **разбиране чрез значения**

към **разбиране чрез корелации**. Машините не знаят *защо* хората действат така, но знаят *как* и *кога* действат — и това им е напълно достатъчно, за да моделират пазара.

Алгоритмична епистемология: знание без разбиране

ИИ системите не работят с класически теории за мотивацията или нуждите (като Maslow, Deci & Ryan и др.). Вместо това те изграждат знание от **емпирична масивност** – от неизмеримо количество поведенчески данни, които не се тълкуват, а се изчисляват.

Този тип знание е *безсмислено* в класическия философски смисъл — то няма нужда от интерпретация. В него е заложена нова форма на рационалност — **корелационна рационалност**, в която значението се ражда от статистическа съвместимост, а не от когнитивно разбиране.

С други думи, ИИ „познава“ човека, без да го „разбира“. Той няма интуиция, морал или емпатия, но има изчислителна точност, която позволява прогнозиране на поведението с вероятност, недостижима за човешката психология.

Поведенчески хипотези и експериментална реалност

Системите на дълбокото обучение (Deep Learning) създават **поведенчески хипотези**, които се проверяват непрекъснато в реално време. Това не е традиционен експериментален метод, а *перманентна емпирика*: всеки потребител е едновременно наблюдателен обект и изследван субект.

Маркетинговите платформи (Meta, Google, TikTok и др.) функционират като **живи лаборатории**, в които поведението на милиарди хора се използва за оптимизация на предсказателните модели.

Така ИИ се превръща в глобална поведенческа инфраструктура, в която:

- всяко действие на потребителя е данна за следващата прогноза;
- всяка прогноза влияе върху следващото действие;
- системата непрекъснато самоусилва своята точност чрез обратна връзка.

Това е **самообучаващ се цикъл на влияние**, при който реалността се адаптира към симулацията, а симулацията се усъвършенства чрез реалността.

Емоции като изчислителни сигнали

В традиционната психология емоцията е вътрешно преживяване. В ерата на ИИ тя се превръща в **алгоритмичен сигнал** — измерима и моделирана величина. Гласовият тон, пулсът, микромимиката или дори начина на писане се превръщат в входни данни за дълбоки невронни мрежи, които разпознават емоционални състояния. Генеративните модели (Generative AI) от своя страна използват тези данни, за да **генерират съдържание, което резонира емоционално**, създавайки илюзия за човешка емпатия.

Например:

- AI чатботове симулират емоционално разбиране, за да създадат доверие в клиентите.
- Рекламни системи предлагат продукти, съобразени с моментното настроение.
- Платформи като Spotify или Netflix адаптират съдържанието според афективните модели на слушателя.

Така **емоцията се превръща в нова валута на пазара**, а *емоционалният профил* — в новата форма на потребителска идентичност.

От влияние към манипулация — ерозията на автономията в алгоритмичната среда

Един от най-острите проблеми в съвременната дигитална икономика е именно **размиването на границата между влияние и манипулация**. Тази граница, макар и тънка, има фундаментално морално значение: *влиянието предполага уважение към автономията*, докато *манипулацията я подменя*.

Влиянието цели да информира, убеди или вдъхнови субекта, като оставя последната дума за решението му. Манипулацията, напротив, **променя самия процес на мислене**, така че решението, което изглежда като свободно, вече е алгоритмично детерминирано.

Алгоритмите като архитекти на поведение

Съвременните маркетингови алгоритми не просто *представят* информация, а я **структурират така, че да насочват вниманието, емоцията и избора**. Те създават това, което Tristan Harris (бивш етичен дизайнер в Google) нарича *“architecture of persuasion”* – архитектура на убеждаването, при която поведението се „подбутва“ в определена посока без съзнателното участие на индивида.

Тази архитектура се изгражда чрез комбинация от когнитивни експлоатации:

- **Ефект на дефицит** (*scarcity effect*): системите изкуствено създават усещане за недостиг („само още 3 бройки“, „времето изтича“).
- **Социално сравнение** (*social proof*): алгоритмите използват поведението на другите като норма („87% от потребителите избраха това“).
- **Алгоритмична пристрастност** (*algorithmic bias*): резултатите и препоръките не са неутрални, а подредени така, че да максимизират ангажираността и консумацията.

Така се изгражда *дизайн на поведение*, в който субектът вече не упражнява рационален избор, а **реагира в рамките на предварително програмирани възможности**.

От свобода на избора към алгоритмична хетерономия

Fenwick (2022) предлага понятието **дигитална хетерономия** – състояние, при което субектът запазва илюзията за избор, но реално се ръководи от предиктивни модели, които изчисляват вероятните му решения. Това е фундаментално нарушение на философския принцип на автономията, формулиран от Кант:

„Автономията е способността на разумното същество да определя собствените си закони.“. В дигиталната среда законите на поведението вече не се определят от разума, а от **алгоритмичната логика на оптимизацията** – логика, чиято цел не е морална, а комерсиална. Потребителят се превръща не просто в наблюдаван обект, а в *управляван процес*, чиято свобода е формална, но не и съществена.

Манипулацията като скрита форма на власт

Традиционните форми на власт (икономическа, политическа, институционална) действат чрез заповеди или санкции. Алгоритмичната власт, напротив, е **невидима и мека** — тя не принуждава, а *предугажда*. Философът Byung-Chul Han (2021) нарича това *“smart power”* — властта, която не се противопоставя на свободата, а я *интегрира* в собствената си логика.

Потребителят чувства, че избира, но реално **реализира избора, който системата е предвидила**. Това е нова форма на *технологически патернализъм* — алгоритмите „знаят по-добре“ кое е добро за теб, защото разполагат с повече информация за твоето поведение, отколкото ти самият.

Архитектурата на вниманието

Shoshana Zuboff (2019) описва икономиката на вниманието като форма на „поведенчески капитализъм“, в която човешкото внимание се превръща в най-ценната стока. Алгоритмите на социалните платформи (YouTube, TikTok, Instagram) са оптимизирани не за истина или качество, а за **задържане на вниманието**.

Всеки елемент на интерфейса – от цвета на бутона до подредбата на съдържанието – е проектиран чрез А/Б тестове, които измерват микрореакции.

Така възниква *архитектура на вниманието*, в която всеки когнитивен стимул се използва като **тригър за поведение**. Потребителят не е просто участник в процеса — той е *вграден в алгоритъма* като източник на данни и като мишена на оптимизация.

Критерий	**Етично влияние**	**Алгоритмична манипулация**
Информираност	Субектът разбира защо му се предлага дадено съдържание.	Съдържанието е персонализирано без знание на субекта.
Свобода на избора	Решението остава в рамките на рационална автономия.	Изборът е предварително моделиран като най-вероятен изход.
Цел на въздействието	Убеждаване чрез информация и аргументи.	Контрол чрез емоционална и когнитивна манипулация.

Манипулацията като инфраструктура

В крайна сметка манипулацията не е просто нежелан страничен ефект на дигиталния маркетинг. Тя е **вградена в неговата архитектура** — в самия начин, по който алгоритмите учат и оптимизират. Тъй като системите се самоусъвършенстват чрез максимизиране на ангажираността, те неизбежно прибягват до експлоатация на когнитивните слабости — това е *функционален резултат*, а не морална грешка.

Fenwick (2022) предупреждава, че ако не бъдат въведени етични стандарти, ще навлезем в епоха на **когнитивен авторитаризъм** — не на насилствена, а на алгоритмична власт, при която контролът се осъществява чрез предсказуемост, а не чрез принуда.

Етични и правни аспекти на алгоритмичния маркетинг

Развитието на изкуствения интелект и дигиталните двойници въвежда безпрецедентна сложност в етичния и правния анализ. Докато традиционните регулаторни режими оперират с концепции като *лични данни*, *информирано съгласие* и *собственост на информация*, съвременните алгоритми функционират на ниво, което надхвърля тези категории. Те **не просто обработват данни**, а *инференциално моделират съзнание* — предвиждат мотивации, емоции и вероятни решения.

Така се появява нова зона на морална и юридическа неяснота: човекът вече не е само обект на наблюдение, а **субект на когнитивно инженерство**.

Алгоритмична непрозрачност — епистемологична и етична криза

Burrell (2019) въвежда понятието за **„algorithmic opacity“** — алгоритмична непрозрачност, при която механизмът на вземане на решения става непознаваем дори за неговите създатели. Това състояние има три нива на непрозрачност:

1. **Техническа непрозрачност** – дълбоките невронни мрежи са толкова комплексни, че процесът на обучение (backpropagation) не подлежи на пряко обяснение.
2. **Институционална непрозрачност** – корпоративните системи умишлено скриват логиката на алгоритмите поради търговски интереси и защита на интелектуална собственост.
3. **Социална непрозрачност** – потребителите не притежават достатъчна дигитална грамотност, за да разберат как и защо се генерират резултатите.

В резултат възниква **епистемологична криза на доверието**: ако дори инженерите не могат да обяснят поведението на системите, на какво основание можем да им се доверим морално или юридически?

Философът Floridi (2023) говори за *„криза на отговорността“*: алгоритмите са автономни в своето вземане на решения, но не и морално отговорни. Така се появява празно пространство — **зона на етична безотговорност**, където последствията са реални, но виновният е неидентифицируем.

Дълбоки фалшификации и експлоатация на идентичността

Появата на *deepfake* технологиите (генеративни модели, способни да синтезират реалистични изображения, видео и глас) бележи нова епоха на **симулация на автентичността**. Тези технологии вече не просто изкривяват реалността, а **пренаписват самата идея за реалност**.

a) Онтологичен разрив между реално и симулирано

Floridi (2023) отбелязва, че deepfake-овете създават „онтологично замърсяване“ – състояние, в което границата между истина и симулация се размива до степен, че доверието губи обективен смисъл. В маркетингов контекст това води до **емоционална експлоатация**: брандове могат да използват дигитални копия на известни личности или „аватари на доверие“, за да повлияят върху нагласите и желанията на потребителите.

Пример: дигитален инфлуенсър, създаден изцяло чрез ИИ, който изглежда като реална личност и „препоръчва“ продукти, без потребителят да знае, че общува със симулация.

b) Етична експлоатация на идентичността

Deepfake технологията отваря въпроса за **собствеността върху личността**. Когато образът, гласът или мимиката на човек могат да бъдат възпроизведени алгоритмично, самата идея за индивидуална идентичност се превръща в дигитален ресурс. Възниква нуждата от ново понятие: *дигитална телесност* — правото на човек да контролира своето дигитално тяло и представяне в алгоритмичната среда.

c) Социално-политически рискове

Deepfake-овете разрушават общественото доверие не само в маркетинга, но и в демократичния дискурс. В среда, където „всичко може да бъде генерирано“, **нищо вече не може да бъде проверено**. Така се създава почва за когнитивен хаос, в който дезинформацията става структурна, а не инцидентна.

Нормативен вакуум — границите на правото в ерата на алгоритмите

Скоростта на технологичните иновации изпреварва способността на правото да реагира. Това води до **нормативен вакуум**, в който основните понятия на правото – субект, воля, отговорност, намерение – се оказват неадекватни за алгоритмичния контекст.

a) GDPR и неговите граници

Общият регламент за защита на данните (GDPR) защитава *личните данни*, но не и *когнитивните структури*, които се извличат от тях. Например, когато алгоритъмът предсказва, че даден потребител е депресивен или импулсивен, това не е „данна“, а *инференция* — следствие от машинно обучение. Следователно, GDPR не защитава съзнанието от **предиктивна интерпретация**.

b) AI Act (2024) — първи, но непълен отговор

Регламентът на ЕС за изкуствения интелект (AI Act) е първият опит да се въведе йерархия на рисковете, но остава **технологично реактивен, а не философски превантивен**. Той разглежда системите според степента им на опасност (минимален, ограничен, висок риск), но не дефинира етичните критерии за когнитивна манипулация.
Няма правна норма, която да забранява системи, способни да влияят върху подсъзнателното поведение.

c) От „нарушение на данни“ към „нарушение на съзнание“

Най-големият пропуск в съвременното право е, че то **защитава информацията, но не и вниманието**. Когато алгоритмите манипулират начина, по който възприемаме реалността, това вече не е кражба на данни, а **кражба на когнитивен капацитет**. Можем да говорим за *epistemic harm* — епистемологична вреда, при която се нарушава способността на човека да различава истина от симулация.

Това изисква ново направление в правната теория: **право на когнитивна неприкосновеност** – защита срещу алгоритми, които въздействат на подсъзнателно ниво и подменят автономията на волята.

Посока на развитие — етика отвъд регулацията

Правото може да санкционира, но не може да възпитава. Следователно решението не е само в регулациите, а в изграждането на **етика на алгоритмичната отговорност**, която включва:

1. **Принцип на алгоритмична прозрачност** – право на обяснение („right to explanation“) при автоматизирани решения.
2. **Принцип на когнитивна справедливост** – забрана за експлоатация на психически уязвимости.
3. **Принцип на цифрово достойнство** – признаване на дигиталната идентичност като неотделима част от личността.

4. **Принцип на технологична умереност** – морален лимит върху степента, в която машините могат да моделират човешко поведение.

Казуси: реални проявления на алгоритмична манипулация

Cambridge Analytica (2018): алгоритмичната манипулация на демокрацията

Фактическа основа

Скандалът с *Cambridge Analytica* избухва през март 2018 г., когато британският вестник *The Guardian* и телевизия *Channel 4* публикуват разследвания, разкриващи, че компанията е събрала данни на **над 87 милиона потребители на Facebook** без тяхното изрично съгласие. Данните са били извлечени чрез приложение за психологическо тестване – *“This Is Your Digital Life”*, създадено от Александър Коган, изследовател от Кеймбриджкия университет. Въпреки че теста е попълнен само от около 270 000 души, Facebook API тогава е позволявал достъп и до данните на техните приятели – практика, която по-късно е забранена.

Методология на манипулацията

Cambridge Analytica комбинира **психографски профили (модел OCEAN)** с алгоритми за *microtargeting* — хиперперсонализирано послание, което използва когнитивни и емоционални слабости. Моделът оценява пет ключови личностни измерения (откритост, съвестност, екстраверсия, съгласие, невротизъм) и на тази основа прогнозира **поведенческа реакция на политически послания**.

Например:

- хора с висока невротичност получавали послания, акцентиращи върху страх от престъпност и имиграция;
- хора с висока откритост – послания за свобода и иновации.

Това превръща Facebook в **невро-политическа лаборатория**, в която емоциите стават инструмент за масово влияние.

Етични и правни измерения

- **Липса на информирано съгласие** — потребителите не са знаели, че данните им ще се използват за политическа агитация.
- **Алгоритмична манипулация на демократичния процес** — посланията са били персонализирани така, че да заобикалят публичния дебат и рационалната дискусия.
- **Регулаторни последствия** — след скандала Facebook беше глобен с **5 млрд. долара** от *Federal Trade Commission (FTC)*, а Cambridge Analytica обяви фалит.
- **Политически ефект** — компанията е работила по кампаниите на *Donald Trump (2016)* и *Brexit Leave EU (2016)*, което повдигна въпроси за ролята на ИИ в подкопаването на изборната автономия.

Deepfake маркетинг (TikTok, 2024): симулацията на автентичността

Фактическа основа

През 2024 г. *Reuters* и *MIT Technology Review* съобщават за нов феномен в социалната мрежа *TikTok*: рекламни кампании, създадени чрез **генеративни ИИ инструменти**, които

използват **реалистични deepfake видеа на известни личности**, без тяхното знание или разрешение.

Един от най-широко разпространените случаи е този с фалшиви реклами на **Елон Мъск, Ким Кардашиян и MrBeast**, които „препоръчват“ криптовалути, инвестиционни схеми или козметика. Някои от тези видеа достигат над 10 милиона гледания преди премахването им.

Механизъм на манипулацията

Технологията за deepfake се базира на **дифузионни генеративни модели** (като *Stable Diffusion* и *DeepFaceLab*), които позволяват с минимален набор от видеоматериали да се генерират изключително реалистични изрази и реч. Рекламните агенции (в някои случаи от сивия пазар на Китай и Източна Европа) създават фалшиви endorsements, за да стимулират доверие и конверсии.

Етични и правни измерения

1. **Нарушение на правото на публичен образ** – deepfake видеата използват лицето и гласа на реални хора без съгласие, което представлява дигитална форма на кражба на идентичност.
2. **Дезинформационен ефект** – потребителите възприемат съдържанието като автентично, особено в кратки формати (Reels, Shorts).
3. **Пазарна манипулация** – рекламите използват фалшив авторитет, което подкопава пазарната прозрачност.
4. **Правна празнина** – повечето юрисдикции все още нямат закони, които директно санкционират създаването на комерсиални deepfake реклами.

Етична последица: разрушаване на доверието

Floridi (2023) нарича това *„пост-истинна етика“* – когато технологиите подменят епистемологичната граница между „реално“ и „създадено“. Deepfake маркетингът подкопава не просто потребителското доверие, а самата идея за **възможност на истина в цифровата комуникация**.

Amazon Recommendation Algorithms: алгоритмична автопойеза на потреблението

Фактическа основа

Amazon използва **рекурентни невронни мрежи и колаборативни филтриращи алгоритми** за своите системи за препоръки още от края на 2000-те. Днес повече от **35% от покупките в Amazon** се генерират чрез препоръчителния механизъм, според данни на *McKinsey & Company (2023)*. Алгоритъмът анализира историята на покупките, времето на разглеждане на продуктите, рейтингите и поведението на други потребители с подобен профил.

Механизъм на самоподдържаща се пристрастност

Този модел създава т.нар. **алгоритмичен затворен цикъл**:

1. Потребителят получава препоръки, основани на предишните му избори.
2. Вероятността да избере нещо извън вече предпочитаните категории намалява.

3. Алгоритъмът приема това като потвърждение и стеснява бъдещите предложения.

Резултатът е **псевдоперсонализация**, при която разнообразието на избори изчезва, а системата създава *предиктивна хомогенизация на потреблението*.

Етични и пазарни измерения

- **Манипулативен дизайн** – алгоритъмът е оптимизиран не за интереса на потребителя, а за максимална конверсия и задържане в платформата.
- **Невидим монопол върху вниманието** – потребителят остава в „екосистемата на Amazon“, без реален стимул да излезе от нея.
- **Пазарна изкривеност** – препоръките подсилват продажбите на продукти, които вече са популярни, което създава ефект на кумулативно предимство („the rich get richer“).
- **Правен аспект** – към момента няма регулация, която да изисква прозрачност за начина, по който препоръчителните системи ранжират продукти.

Казус	Основен проблем	Етичен риск	Правен вакуум
Cambridge Analytica (2018)	Използване на лични данни за психографско профилиране и политическа манипулация	Подкопаване на демократичната автономия	Липса на регулация за психологическо моделиране
Deepfake маркетинг (TikTok, 2024)	Генерирани лица на известни личности за реклами	Симулация на автентичност и измамно доверие	Липса на правен статус на „дигиталната личност“
Amazon Algorithms	Затворен цикъл на препоръки и пристрастия	Ограничаване на избор и когнитивна свобода	Липса на прозрачност в алгоритмичното ранжиране

Заключение

В разгара на дигиталната епоха, където поведението на човека се анализира, предсказва и моделира в реално време, **въпросът за границата между влияние и манипулация** се превръща в основен етичен и философски проблем. Това размиване не е страничен ефект на технологиите, а **структурен синдром на обществото на данните**, което измерва и оптимизира всичко — включително човешката спонтанност.

1. От технологичен детерминизъм към етичен синергизъм

Твърдението, че технологиите неизбежно водят до манипулация, отразява форма на **технологичен детерминизъм** — схващането, че машините управляват развитието на обществото. Но по-задълбоченият анализ показва, че **проблемът не е в алгоритъма, а в неговата ценностна рамка**. Алгоритъмът сам по себе си е неутрален — той е израз на логиката на оптимизация. Етичният въпрос възниква тогава, когато оптимизацията се прилага към човешкото съзнание, без морална цел или граница. Както пише Floridi (2023), *„когато изчислителната рационалност се освободи от моралната рационалност, тя се превръща в инструмент на властта, а не на знанието“*.

Следователно, задачата на етиката в дигиталната икономика не е да спре технологиите, а **да ги вгради в морален синергизъм**, в който алгоритмите не моделират хората, а подпомагат тяхната когнитивна автономия.

2. Свободата в условията на предсказуемост

Класическото разбиране за свобода — като възможност за избор между алтернативи — губи смисъл в свят, в който алгоритмите знаят нашите предпочитания по-добре от самите нас. Поведението ни вече е **обект на статистическа предвидимост**, а не на философска непредвидимост. Това налага ново тълкуване: Свободата в епохата на алгоритмите трябва да се разбира **не като избор**, а като **осъзнаване** — способността да разпознаем, че нашите избори са моделирани. Истинската автономия се измества от равнището на действие към равнището на рефлексия.

Свободата днес не е право да избереш, а умение да осъзнаеш кой избира вместо теб.

Това преосмисляне връща философската мисъл към Кантовото определение на автономията – не като отсъствие на външна принуда, а като **самоналагане на разумен морален закон**. В контекста на дигиталната икономика този морален закон е **етиката на прозрачността и осъзнатостта** – човекът трябва да знае кога и как неговото съзнание е превърнато в инструмент за прогнозиране.

3. Непредсказуемостта като форма на етично съпротивление

В алгоритмична среда, където стойността се създава чрез предсказуемостта на поведението, **непредсказуемостта се превръща в най-висша форма на етична автономия**. Това означава не хаос, а *творческа несъвместимост с машинната логика*.

Както Zuboff (2019) подчертава, поведенческият капитализъм се захранва от „предвидимите остатъци“ на човешкото действие. Следователно, **етичното действие е това, което не може да бъде предсказано от системата** — акт на когнитивна независимост, който избягва алгоритмичния модел.

Тук се появява нова категория в етиката на дигиталната автономия — *епистемологична съпротива*: способността на индивида да защитава правото си да не бъде напълно прозрачен за машините. Това може да се прояви чрез:

- съзнателно нарушаване на алгоритмични шаблони;
- избори, които не следват очакванията на модела;
- отказ от постоянна оптимизация на себе си според данните.

В този смисъл **човешката грешка, непоследователност и ирационалност** придобиват морална стойност — те са доказателство за живото, непредсказуемо съзнание.

4. Етика на съзнанието, а не на данните

Съвременната етика на изкуствения интелект твърде често се концентрира върху защитата на данните (data ethics). Но както показва развитието на алгоритмичните системи, **истинският обект на защита трябва да бъде съзнанието**, а не информацията.

Когато данните за поведението се използват за моделиране на емоции, мотивации и решения, те вече не са просто лични — те са *ментални проекции*.

Етичният въпрос следователно се измества от *„кой притежава данните?“* към *„кой притежава интерпретацията на съзнанието?“*.

Това изисква нова парадигма: **когнитивна етика** — дисциплина, която защитава правото на човека да остане мисловно автономен. Основен принцип на тази етика би бил: „*Алгоритъмът може да познава поведението ми, но не и да определя мисълта ми.“*

5. Отговорността като хоризонт на дигиталното човечество

В епохата на автономни алгоритми отговорността не може да остане само индивидуална. Тя трябва да се мисли **системно** — като колективен морален хоризонт, в който дизайнери, компании и законодатели носят споделена отговорност за когнитивните ефекти на технологиите.

Както отбелязва Helbing (2022), всяка система, която влияе върху общественото поведение, трябва да се подчинява на принципа на *„етична рефлексия в реално време“* - постоянна оценка на въздействието ѝ върху автономията на човека.

Това не е просто регулаторен механизъм, а нов тип **етична култура**, при която човешкото достойнство се мисли не като статична даденост, а като динамична способност да останем *човешки* в контекста на машинното.