



## **ДИГИТАЛНИЯТ НАДЗОР: КРИТИКА НА ПОВЕДЕНЧЕСКОТО МОДЕЛИРАНЕ**

**Билиан Маринов**

[bilinic@tu-sofia.bg](mailto:bilinic@tu-sofia.bg)

*Технически университет – София,  
бул. „Кл. Охридски“ 8,  
БЪЛГАРИЯ*

***Ключови думи:** дигитален паноптикум, дигитален надзор, технологии, етика, техно-тоталитаризъм*

***Резюме:** В статията се изследват ефекти на дигитализацията, които могат да са предизвикателства пред индивидуалната свобода и обществената структура, видяни в обектива на технологичната етика. Въвежда се концепцията „дигитален паноптикум“, осветляваща тенденцията за „програмиране на човека да действа по предписание“, като се дискутират нови форми на социален контрол. Критически се анализира ролята на технологиите в процеса на алгоритмично моделиране на поведението на потребителя и автоматизираното вземане на решения, които могат да подтикнат към дигитални зависимости и ограничаване на автономността. Акценти са „невидимите“ дигитални механизми на глобалния пазарен капитализъм, които предпоставят човешките избори чрез използването на лични данни като икономическа стойност. В тази нова реалност индивидите стават пасивни участници в процеси, които не разбират напълно, като губят контрола върху собствените си избори и желания. Аргументира се тезата, че модерните технологии, като инструментариум за нова икономическа логика, наред с това благоприятстват създаването на тоталитарни модели на подчинение, чрез които властта и информацията се централизират, докато автентичността на изборите, прозрачността в отношенията и свободата на личността са поставени под въпрос. Подчертава се необходимостта от нови регулаторни механизми за постигане на баланс между технологичните иновации и общочовешките ценности, лежащи в основата на демократичните принципи. Сред препоръките за ограничаване на т. нар. „техно-тоталитаризъм“ са: изискване за разбираемост на алгоритмите, децентрализация на дигиталните платформи и етичен дизайн.*

### **ВЪВЕДЕНИЕ**

Дигиталният свят изкушава с илюзията за безгранични възможности, като създава среда, в която всяко човешко действие е наблюдавано и прогнозирано. В тази нова реалност обаче, основният етичен въпрос вече не се изчерпва с това кой контролира технологиите (ТТ), а до това, в каква степен тяхната сила неизбежно води до „централизация на власт“ [1] и намеса в човешката автономност.

ТТ са феномен с двойствена природа (от *pharmakon* – лек и отрова) [2, 3], носители както на обещания за прогрес, така и на потенциал за моделиране на обществени нагласи, междуличностни отношения и индивидуално поведение. Тяхната роля, освен на асистенти–ускорители на човешката мобилност във време и пространство, е и в създаването на нови парадигми за социален ред. Последното се подкрепя от нарастващ брой критици в сферата на технологичната етика. Добре известни са например коментарите на S. Zuboff относно корпорациите, които чрез анализ на лични данни „предсказват и формират“ потребителското поведение като го превръщат в обект на търговия [1]. Чрез дигитализация и скрито наблюдение те целят да реконструират потребителското желание и избори, като мотивират формиране на нов тип зависимости и стандартизирано поведение. В този свят „вече не ние използваме технологиите – те използват нас“ [1]. Тревожни са и опасенията на Y. Narari за влиянието на ТТ върху човека, застрашен от „хакване“, когато дигитални алгоритми разпознават потребностите ни „по-добре от самите нас“ [4]. Те ни водят към манипулирани избори и действия на терена на пазарния капитализъм. Нещо повече, F. Pasquale проблематизира *непрозрачността* при вземане на решения чрез „скрити“ автоматизирани алгоритми, които предотвратяват възможността за лична съпротива. Той заявява: „Когато не знаем как работи една система, не можем да ѝ се противопоставим“ [5]. Така системите за наблюдение, ИИ, биометричният мониторинг и алгоритмичната цензура оформят среда, в която **границата между свобода и контрол е все по-размита**. В нея поведението ни е моделирано, често без нашето съгласие. Заедно с това, алгоритмите, определящи кой да получи работа, кредит или медицинска помощ, оперират със „скрити предразсъдъци“ [6], които могат да увеличат социалните неравенства и разделения.

Направеният по-горе обзор открива важни въпроси за по-широко обсъждане и етични регулации на новите ТТ, които да гарантират **неприкосновеността** на личността: 1) *Кой носи отговорност за възможните рискове, свързани с ТТ?*; 2) *Какви етични задължения имат корпорациите и правителствата, които контролират и използват тези ТТ?*; 3) *Възможно ли е и как постигането на баланс между ефективността и удобството като продукт на ТТ, и творческата същност на човека?*; 4) *Какви етични принципи следва да се прилагат в защита на личните данни в дигиталния свят?*; 5) *Как да избегнем ситуацията, в която ИИ нарушава основни човешки права?*

Целта на изследването е да се установят и по-задълбочено да се осмислят конфликтни области на взаимодействие между човека и ТТ. На основата на критически анализ в технологичната етика и когнитивната психология се аргументира тезата, че, освен че е фактор в трансформацията на човешката комуникация и автоматизацията на процеси, дигитализацията изпълнява функциите на „обществен надзорник“ и „навигатор на индивидуалното поведение“.

## **ОТ КЛАСИЧЕСКИ ТОТАЛИТАРИЗЪМ КЪМ „ДИГИТАЛЕН ПАНОПТИКУМ“**

В исторически план ТТ винаги са използвани за социален контрол – от бюрократичните структури на абсолютните монархии [7] до средствата за пропаганда в тоталитарните режими на XX век [8]. Те радикално трансформират социалната тъкан, правейки индивидите предсказуеми и податливи на влияние и внушения. Технологичната революция обаче доведе до нов тип контрол, наричан „надзорен капитализъм“, който съчетава класически стратегии с дигитални методи на манипулация [1]. Възможността за непрекъснато събиране и анализ на лични данни трансформира самото понятие за контрол – той става невидим, персонализиран и динамичен, като повлиява поведението ни в реално време.

Дигиталните алгоритми често моделират потребителските предпочитания чрез **скрити механизми** [1], които **заобикалят съзнателното и консенсусно вземане на решения**. Социални мрежи, мобилни приложения, алгоритми за персонализирани реклами и системи за препоръки (напр., Netflix, Amazon, YouTube) оформят бъдещето ни още преди да осъзнаем, че сме „програмирани“ да следваме определен път. Тази „метаморфоза на социалната реалност“ в различна степен променя самите условия за формиране на моралното мислене и действие. При това автономността е силно ограничена. Разграничението между видимата и невидимата власт, интерпретирано като модерна форма на тоталитаризъм, е едно от най-ярките прозрения, които се пренасят и в дигиталния свят. Това налага преосмисляне на въпроса за социалния контрол в светлината на класическите идеи за свобода, добродетел и демокрация.

Ако през XX в. тоталитарните режими разчитаха на директен контрол, цензура и репресии, то съвременните ТТ постигат това **по-неуловим начин чрез автоматизирано вземане на решения**. Според F. Pasquale, алгоритмите, които управляват дигиталната среда, работят като „черни кутии“, чиято логика остава скрита дори за регулаторите. Той посочва, че властта вече не се упражнява чрез физическа принуда, а чрез невидими механизми, които оформят възприятията и изборите ни, без дори да го осъзнаваме [5]. Тази идея кореспондира с популярната метафора за „паноптикум“ (J. Bentham), доразвита в структуралистките и постструктуралистките концепции за властта, по-конкретно – от M. Foucault [7]. Той описва паноптикум, където властта се проявява чрез „невидима принуда“, придаваща ѝ ефективност и неоспоримост. Външното упражняване на контрол води до интернализиране на този контрол. Така властта елиминира всяка форма на съпротива чрез оперантно обусловена саморегулация (основен подход на бихейвиоризма). Днес този механизъм е усилен с ИИ и неосезаемото управление на данни [9, 10, 11, 12]. В дигиталната ера подобен принцип се прилага в нов контекст: алгоритмите на ИИ действат като невидими пазители на социалния ред, оформящи нашето поведение и дори представите ни за свободна воля. По този път архитектурата на социалните мрежи и дигиталните платформи продуцира нов вид паноптикум, в който **индивидът не е просто надзираван, а активно програмиран да действа по предписание**. Последното позволява формулиране на понятието „**дигитален паноптикум**“, опериращ по метода на това, което можем да наречем „дигитален надзор“. Неговата функционалност е детерминирана от три основни фактора:

- ◆ **Масово събиране на данни** – корпорации и правителства натрупват безпрецедентно количество информация за всеки потребител;

- ◆ **Алгоритмично вземане на решения** – автоматизираните системи определят кои новини и реклами получаваме и дори какви работни възможности ни се предлагат;

- ◆ **Непрозрачност на технологичния контрол** – ако класическите тоталитарни режими разчитат на принуда, то дигиталните ТТ постигат същия ефект чрез персонализирано съдържание и когнитивни пристрастия [10]. Това кореспондира и с концепцията за **мека намеса** (soft paternalism), която Thaler и Sunstein въвеждат в тяхната теория за „побутването“ (nudge theory), целящо проектиране на избора [12].

Сравнявайки класическите тоталитарни стратегии с дигиталния контрол, можем да заключим, че вторият се отличава със своята **висока ефективност, дискретност и адаптивност**.

Традиционно търсенят от античната и класическа философия баланс между контрол и свобода, в съвременния дигитален свят изглежда нарушен в полза на **централизацията на информация и власт**. Възходът на дигиталния надзор поставя под въпрос традиционните представи за демокрация и автономност. Както твърди В. Bratton дигиталната среда не е просто инструмент за наблюдение – тя пренаписва

правилата на социалните и властовите отношения, създавайки нови форми на зависимост с цикличен характер [13]. Това директно кореспондира с теорията за „**порочните кръгове на допамин**“, развита от А. Lembke [14]. Тя анализира алгоритмите, които стимулират отделянето на допамин, посочвайки, че те го правят непрекъснато, за да ни задържат пред екрана. Така се пренастройва работата на човешкия мозък за повишаване толерантността ни към пристрастяване и намаляване способността за самоконтрол. Зависимостите, както и ефектът на предсказуемост на поведението, са диктувани от потребността за социално взаимодействие, информация и валидиране. Те лесно могат да се култивират в платформи като Facebook, Instagram или Twitter и да станат елемент в **културата на лайковете, автоматизираните реакции и постоянното сравнение с другите**.

Подобни процеси налагат създаването на нова етично-правна регулация и критично преосмисляне ролята на ТТ в обществото. Алгоритмичният контрол може да бъде ефективен и удобен, но без демократичен механизъм за прозрачност и отговорност той носи риск да превърне дигиталното пространство в своеобразен **техно-тоталитаризъм**.

## **АВТОНОМНОСТ И ДЕМОКРАТИЧЕН ДЕФИЦИТ В УСЛОВИЯТА НА ДИГИТАЛЕН ПАНОПТИКУМ**

Поведенческото моделиране в контекста на дигитализация поставя под въпрос основни етични принципи като **автентичността на изборите, свободата на личността, социалната справедливост** и пр. Когато възможността за избор е редуцирана до технологични решения, които не само го предсказват, но го и заместват, е логично да се запитаме: *могат ли дигиталните ТТ да съществуват в синхрон с основните човешки ценности, или неизбежно ще репродуцират форми на доминация, контрол и манипулация?*

Класическата философия разглежда автономността като основополагаща ценност в постигането на морална зрялост и морална отговорност. За Kant последната е символ на идеала за свободна воля, на способността на индивида да се идентифицира не с желанието или външната принуда, а с неговото „по-висше“, рационалното Аз. В дигиталния свят обаче автономността е застрашена от алгоритми, които не просто предлагат избори, а формират условията, при които тези избори се случват. Според Zuboff, когато корпорациите обработват данни за нашите мисли, емоции и поведение, те подменят избора ни с предварително програмирани предразположения [1]. Тук възниква етичната дилема: *може ли да мислим за някого, че е независим, ако неговото поведение е резултат от алгоритмични внушения?*

В съзвучие с горното Е. Mогозов подчертава, че дигитализацията не гарантира демокрация – напротив, тя може да доведе до „авторитаризъм по подразбиране“, ако липсва адекватна регулация [11]. Според него „властта в дигиталната ера принадлежи не на тези, които гласуват, а на тези, които притежават алгоритмите“ [11]. Това поражда дилемата: *може ли демокрацията да функционира в условията на информационна асиметрия, при която корпорации разполагат с несравнимо повече данни и контрол от самите граждани?* В духа на Хабермас, **демокрацията изисква участие в рационален публичен диалог** [15]. Но когато информационният поток е управляван от алгоритми, диалогът става продукт на корпоративни интереси.

Ако приемем, че дигиталният надзор не е неизбежен, то как следва да противодействаме на технологичния тоталитаризъм? От направения анализ дотук, можем да изведем няколко етични правила, които биха възпрепятствали споменатите рискове:

◆ **Прозрачност и отговорност** – създаването на регулаторни рамки, които да изискват разбираемост на алгоритмите [16], получаване на информирано съгласие и поверителност на данните;

◆ **Дигитална децентрализация** – създаване на дигитални платформи с повишена автономност на потребителя, за да се намали концентрацията на власт в ръцете на малък брой технологични корпорации [17];

◆ **Етичен дизайн** – въвеждане на подхода на „технологичната хуманизация“ (включващ емпатия, приобщаване, достъпност), който изисква етично-съобразени алгоритми и поверителност на личните данни.

## ПСИХОЛОГИЧЕСКИ И СОЦИАЛНИ АСПЕКТИ НА ДИГИТАЛИЗАЦИЯТА

Контролът на информацията е централен инструмент за формиране на обществени нагласи и поведение – от класическите техники на пропагандата, описани от E. Bernays (1928) до модерните алгоритмични манипулации [18]. Ако в миналото властта е разчитала на медийна цензура, днес корпорации и политици използват *микротаргетирането*, с цел убеждаване и когнитивно моделиране. Добре проучени са например въздействията на алгоритми в социалните мрежи (СМ) и търсачки (напр., Facebook, Instagram, Google), свързани с явления като „филтърни балони“ и „ехо-камери“ [19, 20]. Въпреки че филтърните балони се формират главно чрез алгоритмично персонализиране на съдържание, а ехо-камерите – в контекста на груповата динамика, и двата феномена водят до **повишена поляризация, ограничаване на критичното мислене и затвърждаване на стереотипни нагласи**, което затруднява диалога и пречатства разширяването на миросгледа. Те създават **изолирана информационна среда**, в която потребителят целенасочено е изложен на само едно конкретно съдържание, което да потвърди неговите предпочитания, и така игнорира или не получава нова и различна от неговата перспектива.

Освен във формирането на нагласи, ТТ участват и в моделиране на емоции и реакции на масите. Нобеловият лауреат психологът D. Kahneman, в изследването си върху **когнитивните пристрастия**, подчертава, че хората са силно податливи на автоматизирани решения, които изискват минимални когнитивни усилия [10]. Това е ключов механизъм, успешно прилаган в СМ. Състои се от алгоритми, които *задържат вниманието* на потребителя като предизвикват реакции (възмущение, радост и др.), без да осъзнаваме тази динамика. Този феномен е известен като „**икономика на вниманието**“. В тази връзка Tristan Harris, бивш дизайнер в Google и критик на дигиталния маркетинг, описва как платформи като Facebook, Instagram и Twitter използват този механизъм, за да *максимизират времето, което потребителите прекарват онлайн* (вж. филма “The Social Dilemma”). Той обяснява, че алгоритмите, проектирани по този модел, експлоатират „слабости“ на човешката психика, засилвайки тревожността, социалния натиск и емоционалната нестабилност. В други проучванията (напр., Twenge) сред по-млади хора, въздействието на СМ включва още увеличаване на зависимостите, депресия и социална изолация [21].

Дигитализацията не само променя когнитивни процеси, но и преобразува поведението на масите. От класическите анализи на G. Le Bon и Z. Freud върху психологията на тълпите ни е известно, че **масите действат ирационално, водени от емоции и внушения**. В дигиталната среда тази динамика се усилва експоненциално, тъй като СМ импулсират групови поляризации, водейки до по-бързо разпространение на крайни възгледи. Алгоритмите не просто предоставят информация, а активно насочват потребителите към по-радикални позиции, защото те задържат повече внимание и ангажираност. Те не различават истина и лъжа, умереност и екстремност – те просто оптимизират процеси за максимално взаимодействие [22]. Това би довело до

създаване на общества, в които дезинформацията, конспиративни теории и политически манипулации се разпространяват по-ефективно от всякога.

От така направените проучвания и анализ на данни, можем да посочим три по-важни рискови аспекта на влияние на дигиталните алгоритми върху индивида и масите:

◆ **Загуба на критично мислене** поради постоянния приток на персонализирана информация, която не подлежи на съмнение или проверка;

◆ **Засилване на социалната поляризация**, тъй като алгоритмите насърчават радикални мнения и емоционални реакции;

◆ **Емоционално и когнитивно изтощение**, причинено от свръхстимулация и постоянна онлайн ангажираност на потребителите в дигитални платформи.

## ЗАКЛЮЧЕНИЕ

Съвременните технологии могат не само да профилират индивида в глобалната мрежа – те моделират неговото поведение, вземат решения вместо него и създават **незабележими, но мощни форми на социален контрол**. Ако в миналото някои държави налагаха тоталитаризъм чрез репресия, днес алгоритмите могат да постигат сходни ефекти чрез персонализирана информация и поведенческа икономика.

Неизбежна и социално потребна е разумна и приемлива степен на дигитален надзор, т.е.: съвременни регулаторни подходи, прозрачност и критичен анализ на дигиталната власт, като неизбежно условие за възпиране на манипулации и злоупотреби чрез поведенческо моделиране, което поставя фундаментални етични въпроси за свободата и автономността на индивида.

## ЛИТЕРАТУРА:

- [1] Zuboff, S. The Age of Surveillance Capitalism. PublicAffairs, ISBN-10:610395697, 2019.
- [2] Derrida, J. Plato's Pharmacy. In: Dissemination, University of Chicago Press, 1981.
- [3] Stiegler, B. What makes life worth living: on pharmacology. UK: Polity, 2013.
- [4] Harari, Y. Homo Deus: A Brief History of Tomorrow. Harper, ISBN-10:0062464310, 2017.
- [5] Pasquale, F. The Black Box Society. Harvard Univ. Press, ISBN-10:0674970845, 2016.
- [6] O'Neil, C. Weapons of Math Destruction. Crown, ISBN-10: 0553418815, 2016.
- [7] Foucault, M. Discipline and Punish. Vintage Books, ISBN: 0679752552, 1977.
- [8] Arendt, H. The Origins of Totalitarianism. Meridian Books, Cleveland & N.Y., 1951.
- [9] Brunton, F. & Nissenbaum, H. Obfuscation. MIT Press. ISBN: 9780262331319, 2015.
- [10] Kahneman, D. Thinking, Fast and Slow. Farrar, Straus & Giroux, N. Y., 2011.
- [11] Morozov, E. The Net Delusion. PublicAffairs, ISBN-13: 978-1586488741, 2011.
- [12] Thaler, R. & Sunstein, C. Nudge. Penguin Books, ISBN: 9780143137009, 2021.
- [13] Bratton, B. The Stack. The MIT Press, England, ISBN: 978-0-262-02957-5, 2015.
- [14] Lembke, A. Dopamine Nation. Dutton, ISBN-13: 978-1524746728, 2021.
- [15] Habermas, J. The Structural Transformation of the Public Sphere. The MIT Press, 1991.
- [16] Pasquale, F. New Laws of Robotics. Harvard Univ. Press, ISBN 9780674975224, 2020.
- [17] Lanier, J. You Are Not a Gadget: A Manifesto. Vintage, ISBN-10: 0307389979, 2011.
- [18] Bernays, E. Propaganda. Horace Liveright, N.Y., 1928.
- [19] Sunstein, C. #Republic. Princeton University Press, ISBN-10: 0691175519, 2017.
- [20] Pariser, E. The Filter Bubble. Penguin Books, ISBN-10: 0143121235, 2012.
- [21] Twenge, J. iGen. Atria Books, ISBN-10: 9781501151989, 2017.
- [22] Tufekci, Z. Twitter and Tear Gas. Yale Univ. Press, ISBN 978-0-300-21512-0, 2017.

# DIGITAL SURVEILLANCE: A CRITIQUE OF BEHAVIORAL MODELING

**Bilian Marinov**

[bilinic@tu-sofia.bg](mailto:bilinic@tu-sofia.bg)

*Technical University of Sofia  
8 Kliment Ohridski Blvd  
BULGARIA*

**Key words:** *digital panopticon, digital surveillance, technology, ethics, techno-totalitarianism*

**Abstract:** *This article examines the effects of digitalization that may pose challenges to individual freedom and societal structure, viewed through the lens of technological ethics. The concept of the "digital panopticon" is introduced, shedding light on the tendency to "programming the human to act by prescription" while discussing new forms of social control. The role of technologies in the process of algorithmic modeling of user behavior and automated decision-making is critically analyzed, with an emphasis on how these processes may push individuals toward digital dependencies and limit autonomy. The focus is on the "invisible" digital mechanisms of global market capitalism, which precondition human choices by utilizing personal data as an economic asset. In this new reality, individuals become passive participants in processes they do not fully understand, losing control over their own choices and desires. The thesis argues that modern technologies, as instruments of a new economic logic, also facilitate the creation of totalitarian models of subjugation, through which power and information are centralized, while the authenticity of choices, transparency in relationships, and individual freedom are called in question. The need for new regulatory mechanisms is emphasized, in order to achieve a balance between technological innovations and universal human values that underpin democratic principles. Among the recommendations to limit so-called "techno-totalitarianism" are: the requirement for algorithm transparency, decentralization of digital platforms, and ethical design.*