



SENTRIX

RETHINKING SENTENCING
THROUGH CHOICE
ARCHITECTURE

Vedenjska ekonomija in pravo
Podjetje in delo 6-7/2025/LI

**DIGITALNI DREGLJAJI IN TEMNI
VZORCI – KAKO TEHNOLOGIJA
IZRABLJA NAŠE VEDENJSKE
ŠIBKOSTI**

Manja Skočir

Building a decision-making matrix

Manja Skočir

Digitalni dregljaji in temni vzorci – kako tehnologija izrablja naše vedenjske šibkosti

1. Uvod
2. Od analognih do digitalnih dregljajev
3. Tipologija digitalnih dregljajev
 - 3.1. Generični digitalni dregljaji
 - 3.2. Personalizirani in adaptivni digitalni dregljaji
4. Temni vzorci
5. Normativna analiza: etični ugovori in pravni odzivi
 - 5.1. Avtonomija
 - 5.2. Svoboda izbire
 - 5.3. Manipulacija
 - 5.4. (Ne)transparentnost
6. Pravni odziv
7. Sklep

Povzetek

Članek analizira pojav digitalnih dregljajev in temnih vzorcev kot enega ključnih načinov, prek katerih sodobne tehnologije izrabljajo kognitivne ranljivosti uporabnikov. Teoretični okvir vedenjske ekonomije – zlasti teorija dregljajev in koncept izbirne arhitekture – omogoča pojasnitev mehanizmov vplivanja, vgrajenih v digitalne tehnologije, ter normativno presojo njihove dopustnosti. Ker se ljudje pogosto odločamo hitro, intuitivno in pristransko, smo posebej dovzetni za subtilne posege v kontekst odločanja. V digitalnem okolju se ti posegi krepijo zaradi lastnosti tehnološke infrastrukture (stalna prisotnost, podatkovna analitika, personalizacija), kar vodi v pojav generičnih, personaliziranih in adaptivnih dregljajev, ki se v skrajni obliki prelijejo v temne vzorce – namerno oblikovane vmesnike, ki zavajajo uporabnike in jih usmerjajo v vedenja proti njihovim interesom. Normativna analiza pokaže, da se klasični ugovori proti analognim dregljajem (ogrožanje avtonomije, omejevanje svobode izbire, manipulativnost, netransparentnost) v digitalnem okolju še zaostrijo. Pravno gledano okvir EU sicer vzpostavlja zaščito le pred temnimi vzorci (GDPR, Direktiva o nepoštenih praksah, DSA, Akt o UI), medtem ko digitalni dregljaji kot subtilnejše oblike vplivanja ostajajo v sivi coni med dopustnim prepričevanjem in prepovedano manipulacijo.

Ključne besede: digitalni dregljaji, temni vzorci, spletna manipulacija, vedenjska ekonomija

Manja SKOČIR, magistrica prava, magistrica filozofije in primerjalne književnosti, mlada raziskovalka na Inštitutu za kriminologijo pri Pravni fakulteti v Ljubljani in doktorska študentka na Pravni fakulteti Univerze v Ljubljani, e-naslov: manja.skocir@inst-krim.si

Abstract

This article analyzes the phenomenon of digital nudges and dark patterns as key ways in which contemporary technologies exploit users' behavioral vulnerabilities. The theoretical framework of behavioral economics—particularly nudge theory and the concept of choice architecture—makes it possible to explain the mechanisms of influence embedded in digital technologies and to assess their normative acceptability. Because people often decide quickly, intuitively, and with cognitive biases, we are especially susceptible to subtle interventions in the decision-making context. In digital environments, these interventions are amplified by the characteristics of technological infrastructure (constant presence, data analytics, personalization), leading to the emergence of generic, personalized, and adaptive nudges, which in their extreme form turn into dark patterns—deliberately designed interfaces that mislead users and steer them toward behaviors contrary to their interests. Normative analysis shows that the classical objections to analog nudges (threats to autonomy, restriction of freedom of choice, manipulateness, non-transparency) become even more pronounced in digital contexts. From a legal perspective, the EU framework provides some protection against dark patterns (GDPR, Unfair Commercial Practices Directive, DSA, AI Act), while digital nudges, as more subtle forms of behavioral influence, remain in a gray area between permissible persuasion and prohibited manipulation.

Keywords: digital nudges, dark patterns, online manipulation, behavioral economics

1. Uvod

Vedenjska ekonomija je v zadnjih desetletjih temeljito spodkopala idealiziran model racionalnega posameznika, ki se v raznolikih in nepredvidljivih življenjskih okoliščinah odloča v skladu s svojimi stabilnimi preferencami in načelom maksimiranja lastnih koristi. Empirične raziskave, ki jih od sedemdesetih let prejšnjega stoletja pod okriljem te interdisciplinarne panoge pospešeno izvajajo psihologi, ekonomisti in kognitivni znanstveniki, so pokazale, da človeško odločanje pogosto odstopa od aksiomov racionalnosti, kakršne predpostavlja klasična ekonomija.¹ V nasprotju z njimi nam vedenjska ekonomija slika podoba »omejeno racionalnih«² ljudi, ki s(m)o podvrženi tako uporabi številnih – običajno sicer koristnih – kognitivnih bližnjic oziroma heuristik³ kot tudi vplivu čustev in strasti.⁴ Drugače kot na homo economicusa je na takšnega posameznika mogoče učinkovito vplivati že z majhnimi, včasih skoraj banalnimi spremembami okolja, v katerem se odloča,⁵ in z uporabo t. i. univerzalnih orožij prepričevanja, ki temeljijo na globoko zakoreninjenih socialnih in kognitivnih vzorcih.⁶

Ta – tudi s prejeto Nobelovo nagrado pripoznana – spoznanja so po eni strani omogočila bolj realistično razumevanje človeškega vedenja, po drugi pa so odprla vrata novim možnostim instrumentalizacije ljudi, saj je ustaljene in evolucijsko pogojene kognitivne mehanizme mogoče uporabiti (tudi) kot orodje za manipulacijo. Vedenjska ekonomija torej ni le postavlila diagnoz in ponudila nekaterih »zdravil«⁷ zanje, temveč je, morda nehote, odprla tudi arze-

¹ A. Tversky, D. Kahneman (1974; 1981); R. H. Thaler (2019), str. 19–113.

² G. Wheeler (2024).

³ D. Kahneman (2017), str. 151–273.

⁴ A. Damasio (2009); D. Ariely (2010).

⁵ A. Tversky, D. Kahneman (1981); R. H. Thaler, C. R. Sunstein (2022).

⁶ R. Cialdini (2015), str. 16–35.

nal učinkovitih orodij za instrumentalizacijo odločevalcev.⁷ Samodejni vzorci, ki aktivirajo intuitivno razmišljanje in delovanje ter nam omogočajo, da se znajdemo v kompleksnem in hitro spreminjajočem se svetu, nas pač delajo zelo ranljive pred vsakim, ki pozna in razume logiko njihovega delovanja ter jo je voljan izkoristiti, da bi vplival na naše odločitve. Enega najopaznejših primerov takšne instrumentalizirajoče rabe spoznanj vedenjske ekonomije predstavljajo elementi zasnove sodobnih digitalnih tehnologij, s katerimi se ukvarjam v tem besedilu. Njihovi oblikovalci se pri strukturiranju uporabniških vmesnikov opirajo na kognitivne pristranskosti in hevristike, identificirane v vedenjskih znanostih, ter jih uporabljajo za usmerjanje pozornosti uporabnikov in subtilno oblikovanje njihovih odločitev.⁸

V tem prispevku vedenjsko ekonomijo – zlasti teorijo dregljajev in koncept izbirne arhitekture – uporabim kot teoretično izhodišče za dvoje. Prvič, za pojasnitev manipulativnih mehanizmov vplivanja na vedenje, ki so del zasnove sodobnih digitalnih tehnologij. Drugič, za normativno presojo dopustnosti takšnega vplivanja: etične ugovore, izrečene proti analognim dregljajem (ogrožanje avtonomije, prikrita manipulativnost, netransparentnost, instrumentalizacija), je namreč mogoče aplicirati tudi na digitalno okolje, pri čemer se pokaže, da se ti pomembno zaostrijo. V članku razvijam tipologijo treh vrst digitalnih dregljajev – generičnih, personaliziranih in adaptivnih. Pokažem, da je dregljaje, tudi digitalne, kljub njihovi etični ranljivosti mogoče ob določenih pogojih razumeti kot sprejemljivo obliko vedenjskega vplivanja. Povsem drugače pa velja za temne vzorce, ki sicer delujejo na podlagi podobnih kognitivnih mehanizmov kot dregljaji (izkoriščanje pristranskosti, hevristik in avtomatiziranih odločitev), vendar se to izvaja na sistematično netransparenten, zavajajoč in instrumentalizirajoč način. Kot takšni pomenijo kvalitativni preskok iz območja še dopustnega vedenjskega vplivanja v sfero nedopustne manipulacije. Prispevek sklenem s kratkim pregledom veljavne pravne ureditve ter presojo, ali ta zagotavlja zadostno varstvo pred manipulativnimi praksami sodobnih digitalnih tehnologij.

2. Od analognih do digitalnih dregljajev

V klasični ekonomiji in tudi v nekaterih drugih družboslovnih znanostih sta dolgo veljali dve hipotezi o človeški naravi. Prva je trdila, da smo ljudje racionalna bitja, ki praviloma sprejemamo logične in preišljene odločitve. Druga je odstopanja od takšnih odločitev razlagala predvsem z vplivom psihičnih stanj, kakršna so strah, naklonjenost ali sovraštvo.⁹ Na teh predpostavkah je zrasla teorija pričakovane koristi, v skladu s katero z veličastnim kognitivnim aparatom in jasno prioritizacijo preferenc opremljen posameznik pri izbiri med alternativnimi možnostmi vselej izbere tisto, ki mu prinaša največjo korist, saj je sposoben pravilno oceniti verjetnost negotovih dogodkov, njegove preference pa so neodvisne od tega, kako so možnosti predstavljene.¹⁰

Čeprav so kritike tega modela obstajale že prej,¹¹ se zdi, da so teorijo pričakovane koristi spodnesle šele raziskave s področja vedenjske ekonomije, ki so pokazale, da gre za teorijo, ki ne opisuje, kako se ljudje v resnici odločajo, ampak kako bi se morali racionalno odloča-

⁷ R. Rebonato (2012), str. 118–122.

⁸ D. Susser *et al.* (2019)

⁹ D. Kahneman (2017).

¹⁰ T. Strle, O. Markič (2021), str. 21–24; R. A. Briggs (2023).

¹¹ H. A. Simon (1957), str. 24–25; P. Slovic (1987).

ti. Ključen očitek teoriji pričakovane koristi, ki ga izreka vedenjska ekonomija, je, da gre za prikrito normativno in ne deskriptivno teorijo, zaradi česar ne more ustrezno pojasniti dejanskih odločitev posameznikov, ampak lahko postavi zgolj njihove norme oziroma ideale. Zasluga vedenjske ekonomije torej ni toliko v tem, da bi razkrila nekaj povsem novega, ampak v tem, da je to, kar je desetletja prej naznanil že Freud, da namreč človek »ni gospodar niti v lastni hiši«, eksperimentalno potrdila, kvantificirala in umestila v ekonomsko vedo.

Vedenjska ekonomija pa ni zgolj razkrila ranljivosti in pomanjkljivosti človeškega odločanja, ampak je ponudila tudi nekatere pristope za njihovo obvladovanje: kognitivne pristanske so sicer vir slabih odločitev, vendar jih je – ob ustrezni zasnovi okolja odločanja – mogoče usmeriti v korist odločevalca. Richard H. Thaler in Cass R. Sunstein sta znamenito ugotovitev Amosa Tverskega in Daniela Kahnemana, da namreč preference ljudi niso stabilne, saj na to, katero izmed alternativnih možnosti bomo izbrali, pomembno vpliva okvir, v katerem so predstavljene,¹² nadgradila z razvojem tehnike vedenjskega vplivanja, imenovane dregljaji (angl. *nudge*). To teorijo, sprva predstavljeno v znanstvenih člankih,¹³ sta Thaler in Sunstein širši javnosti približala z monografijo Nudge (2008),¹⁴ ki je sprožila val zanimanja za dregljaje kot neinvazivno alternativo klasični regulaciji.

Dregljaji temeljijo na predpostavki, da se ljudje v vsakdanjih situacijah navadno odločamo hitro, avtomatično in brez poglobljenega razmisleka, saj pri tem pogosto (nevede) uporabljamo avtomatizirane miselne bližnjice oziroma hevrstike, ki pa lahko vodijo v sistematične napake v presoji. Denimo: v naboru različnih izbir prvo navedeno možnost običajno dojemamo kot privzeto oziroma kot najoptimalnejšo odločitev. Ker sledenje takšni možnosti zmanjšuje potrebo po razmisleku (saj jo hkrati razumemo kot implicitno priporočilo), lahko na izbire ljudi pomembno vplivamo že s tem, da na primer okolju prijaznejšo možnost postavimo na začetek nabora. Podobno je na izbiranje ljudi mogoče vplivati s t. i. sidranjem cen; prečrtana prvotna cena ob navedbi nove, znižane, ustvarja vtis ugodnejše ponudbe ter deluje kot močan spodbujevalec k nakupu.

Številne zasebne in javne ustanove so začele uvajati vedenjskoekonomске pristope, saj se je zdelo, da je z dregljaji mogoče pomembno izboljšati odločitve ljudi glede zdravja, finančnega položaja in splošnega blagostanja, ne da bi s tem posegali v njihovo svobodo izbire.¹⁵ Toda kljub navdušenju so dregljaji sprožili tudi številne pomisleke in odprli razprave o etičnosti tovrstne regulacije.¹⁶ Kritiki poudarjajo, da dregljaji ogrožajo avtonomijo, vsebujejo elemente manipulacije, omogočajo zasledovanje nelegitimnih ciljev, so netransparentni in temeljijo na paternalističnih predpostavkah o interesih posameznika. Poleg tega lahko prispevajo k depolitizaciji družbenih vprašanj, saj odgovornost za reševanje družbenih problemov preusmerjajo z institucij na posameznike.¹⁷

Vendar je uporaba dregljajev presegla analogne kontekste, še preden so bile etične in normativne razprave o njihovi legitimnosti temeljito razdelane in evaluirane. S pojavom in

¹² A. Tversky, D. Kahneman (1981), str. 455–458.

¹³ R. H. Thaler, C. R. Sunstein (2003a; 2003b).

¹⁴ R. H. Thaler, C. R. Sunstein (2008).

¹⁵ Za pregled različnih intervencij vedenjskih enot po svetu glej OECD (2017), za razvoj v Evropi pa poročilo Evropske komisije (2016).

¹⁶ A. Schmidt, B. Engelen (2020).

¹⁷ *Ibidem*; T. Grüne-Yanoff (2012), R. Rebonato (2012).

krepitvijo digitalizacije v vsakdanjem življenju so se namreč spoznanja vedenjske ekonomije začela pospešeno uporabljati tudi pri oblikovanju digitalnih okolij. Ugotovitve vedenjskih znanosti o odločanju posameznikov so se izkazale kot izjemno učinkovita orodja za uresničevanje različnih ciljev: od izboljševanja uporabniške izkušnje in s tem podaljševanja časa in pozornosti, ki ju uporabniki namenjajo digitalnemu okolju, do varovanja njihovih interesov (na primer z usmerjanjem k izbiri varnejših gesel) ter subtilnega oblikovanja potrošniških in političnih izbir.

Za nabor praks, ki koreninijo v spoznanjih vedenjske ekonomije in prek oblikovanja digitalnega okolja poskušajo (bolj ali manj prikrito) vplivati na vedenje in odločanje posameznikov, so teoretiki začeli uporabljati izraz »digitalni dregljaji« (angl. *digital nudging*).¹⁸ Podobno kot v analognem kontekstu je bila tudi v digitalnem ideja dreglajev sprva predstavljena kot dobrodošla oblika regulacije vedenja, kar potrjujejo zgodnje znanstvene objave s tega področja. Njihovi avtorji so v digitalnih dregljajih prepoznavali orodja, s katerimi bi bilo mogoče po eni strani okrepiti obstoječe analogne vedenjske intervencije, po drugi strani pa spodbujati »boljše« odločanje uporabnikov v zvezi s specifičnimi izbirami v digitalnem okolju – denimo pri krepitvi kibernetske varnosti in zaščiti zasebnosti.¹⁹

Kmalu se je razkrilo, da je z digitalnimi dregljaji poleg krepitve blaginje uporabnikov mogoče večati tudi dobičke podjetij. Še več: očitno je postalo, da digitalni dregljaji niso samo preslikava analognih pristopov v novo okolje, temveč so kvalitativno drugačen pojav – so namreč rezultat prepleta vedenjskih znanosti in digitalne tehnološke infrastrukture, katere značilnosti pomembno spreminjajo moč, domet, pogostost in intenzivnost vedenjskih posegov. Čeprav je mogoče analogne dregljaje – kljub resnim etičnim pomislekom – razumeti kot dopustno obliko regulacije, če so transparentni, reverzibilni in usmerjeni v interes posameznika, digitalni dregljaji pomenijo zasuk, ki zahteva dodatno presojo njihove etične upravičenosti. Zlasti tisti dregljaji, ki so se v digitalnem okolju spojili z velikim podatkovjem (angl. *big data*), so se izkazali ne le kot bistveno učinkovitejši, ampak tudi etično bolj tvegani od svojih analognih predhodnikov.²⁰

3. Tipologija digitalnih dreglajev

Teorija dreglajev je raziskovalcem informacijskih sistemov in digitalnih tehnologij ponudila nov konceptualni aparat tako za oblikovanje kot tudi za razumevanje vplivov na vedenje v digitalnih okoljih. Pri opredelitvi digitalnih dreglajev se referenčni avtorji opirajo na izvorno definicijo dreglajev, v skladu s katero je dregljaj »katerikoli vidik izbirne arhitekture, ki na predvidljiv način spreminja vedenje ljudi, pri tem pa ne prepoveduje nobene možnosti niti bistveno ne spreminja njihovih finančnih spodbud«.²¹ Digitalni dregljaj je tako mogoče opredeliti kot namenjen poseg v digitalnem ali hibridnem (angl. *blended*) okolju, katerega namen je na predvidljiv način vplivati na posameznikovo presojo, izbiro ali vedenje z uporabo spoznanj o kognitivnih omejitvah, pristranskostih ter utrjenih rutinah in navadah v odločanju.

¹⁸ D. Hummel *et al.* (2018); C. Meske, T. Potthoff (2017); T. B. Lembcke (2019); M. Weinmann *et al.* (2016).

¹⁹ M. Weinmann *et al.* (2016).

²⁰ T. B. Lembcke (2019), str. 9–13.

²¹ R. H. Thaler, C. R. Sunstein (2022), str. 26.

Čeprav se je koncept dregljaja skoraj neposredno prevedel iz analognega v digitalno okolje, to še ne pomeni, da se je hkrati prenesla tudi njegova etična obravnava. V definicijskem smislu digitalni dregljaj še vedno pomeni vidik izbirne arhitekture, ki na predvidljiv način spreminja vedenje ljudi, ne da bi pri tem prepovedoval možnosti ali bistveno spreminjal njihove finančne spodbude. Toda v nadaljevanju bom pokazala, da je izpolnjevanje pogojev, ki analogne dregljaje lahko utemeljijo kot legitimno obliko regulacije (transparentnost, reverzibilnost, usmerjenost v interes posameznika), v digitalnem okolju bistveno težje. Prav zato zahteva digitalno okolje lastno tipologijo, ki omogoča razlikovanje med različnimi oblikami digitalnih dregljajev in oceno njihove etične dopustnosti.

Analiza digitalnih dregljajev pokaže, da jih je mogoče razvrstiti v vsaj dve osnovni skupini. Prvo sestavljajo neposredni nasledniki analognih dregljajev, ki so namenjeni »povprečnemu uporabniku« in delujejo na podlagi načela »ena rešitev za vse«. Veliko učinkovitejši – a tudi etično bolj problematični – so personalizirani digitalni dregljaji, ki jih analogno okolje ne dopušča, saj temeljijo na analizi uporabniških podatkov ter svojo moč črpajo iz natančne prilagojenosti interesom, značilnostim in (celo) trenutnemu razpoloženju posameznika. Kot bom razložila v nadaljevanju, se takšne intervencije bistveno odmikajo od izvirne ideje dreganja in približujejo praksam, t. i. temnim vzorcem, ki jih literatura prepozna kot nelegitimne oblike manipulacije in ki jih pravo EU v določenih primerih že prepoveduje.

3.1. Generični digitalni dregljaji

Generični digitalni dregljaji so tako kot njihovi analogni predhodniki neindividualizirani. Njihov cilj je vplivati na povprečnega posameznika, glede katerega so eksperimenti vedenjskih znanstvenikov pokazali, da je pretirano samozavesten, preveč navezan na že pridobljene dobrine, nagnjen k sprejemanju privzetih izbir, podvržen vplivu prvega vtisa, šibak v statističnem razmišljanju ter čustveno nestanovit in dovzeten za socialni pritisk. Vendar selitev v digitalno okolje ne pomeni zgolj prenosa analognih pristopov vanj, temveč njihovo pomembno preoblikovanje.²²

Digitalno okolje ima namreč določene specifične: empirične raziskave so pokazale, da krepi intuitivno odločanje oziroma mehanizme »sistema 1«, saj bolj poudarja vizualne dražljaje.²³ Poleg tega so digitalna okolja pogosto podatkovno zasičena, kar vodi v informacijsko preobremenjenost in posledično zmanjšuje pozornost ter zmožnost poglobljenega premisleka.²⁴ Temu pritrjuje tudi avtor danes že skoraj kultne knjige Plitvine Nicholas Carr, ki kot eno ključnih posledic izpostavljenosti digitalnemu okolju za človeške možgane vidi upad sposobnosti kritičnega razmisleka. Carr pojasnjuje, da je kritično razmišljanje med brskanjem po spletu sicer mogoče – tako kot je mogoče tudi plitko razmišljanje med branjem knjige –, vendar digitalna tehnologija takšnega načina razmišljanja ne spodbuja in ne nagrajuje, ampak (nasprotno): vse bolj odrija linearno razmišljanje in ga nadomešča s t. i. mrežnim razmišljanjem.²⁵

O digitalnem okolju je torej smiselno razmišljati kot o okolju, ki šibi sposobnost koncentracije in dela posameznike še ranljivejše glede napak pri presoji in odločanju. Poleg krepitve pristranskosti, značilnih tudi za analogna okolja, so raziskovalci v digitalnem okolju iden-

²² D. Kahneman (2017), str. 29–273.

²³ S. Benartzi, J. Lehrer (2015).

²⁴ L. Ziming (2005).

²⁵ N. Carr (2010), str. 45–70.

tificirali nove, specifične oblike pristranskosti – denimo *middle bias*, tj. težnjo uporabnikov, da pogosteje izberejo možnosti, umeščene na sredino zaslona.²⁶ Zato generični digitalni dregljaji niso zgolj preslikava analognih, temveč njihova nadgrajena različica, ki zaradi posebnosti digitalnega okolja razpolaga z večjo močjo in prožnostjo. Poleg tega je digitalno okolje bistveno lažje in ceneje prilagodljivo od analognega. Medtem ko so posegi v fizična okolja pogosto dragi in zapleteni (na primer spremembe v zakonodaji ali infrastrukturi), lahko oblikovalci digitalnih okolij zasnovno spletne strani ali aplikacij spreminjajo hitro in z minimalnimi stroški ter sproti testirajo, kateri pristopi so najučinkovitejši.

3.2. Personalizirani in adaptivni digitalni dregljaji

V digitalnem okolju se ideja prilagajanja izbirne arhitekture ni spojila le s specifičnimi lastnostmi tega okolja, ampak tudi z velikimi podatkovnimi zbirkami (angl. *big data*). Z podatkovno analitiko je podatke uporabnikov spletnih storitev mogoče pretapljati v – ekonomsko zelo dobičkonosno – zbirko osebnostih in značajskih lastnosti posameznikov. Na tej osnovi je uporabnike mogoče profilirati in prek njihovim lastnostim prilagojene spletne izbirne arhitekture vplivati na njihovo razpoloženje, vedenje ter (potrošniške in politične) izbire.²⁷ Intervencije v arhitekturo odločanja, ki združujejo spoznanja vedenjske ekonomije, značilnosti digitalnega okolja in podatkovno analitiko, bom imenovala personalizirani digitalni dregljaji.

Še korak dlje gredo dregljaji, ki se v realnem času prilagajajo vsakokratnim odzivom in razpoloženju uporabnika, zato jih imenujem adaptivni digitalni dregljaji. Medtem ko personalizirani dregljaji temeljijo na profilu, zgrajenem na podlagi uporabnikove spletne zgodovine, se adaptivni sproti spreminjajo na osnovi tekoče analize podatkov in omogočajo stalno prilagajanje izbirne arhitekture.

Takšna sprememba – od pavšalnih analognih dregljajev do personaliziranih in adaptivnih – sama po sebi še ne pomeni nujno prestopa v sfero nelegitimnega in (kot bomo videli) pravno prepovedanega. Vsekakor pa pomembno povečuje tveganja, da se vedenjski vplivi spremenijo v sistematično manipulacijo, kar jih približuje praksam, ki jih bom v nadaljevanju obravnavala kot temne vzorce.

4. Temni vzorci

Posebno raven vedenjskega vplivanja v digitalnem okolju pomenijo temni vzorci (*dark patterns*). To je pojav, ki ga lahko razumemo kot radikalizacijo digitalnega dreganja, vendar zaradi svoje izrazite manipulativnosti in sistematičnega izigravanja uporabnikov zahteva samostojno obravnavo.

Pojem temni vzorci je leta 2010 skoval Harry Brignull, načrtovalec uporabniške izkušnje (angl. *user experience designer*), v zadnjem desetletju pa se je to poimenovanje uveljavilo v akademskih krogih, množičnih medijih in tudi v zakonodajnih predpisih.²⁸ Čeprav v referenčni literaturi obstaja več definicij tega pojava, ga v tem prispevku razumem kot premišljeno zasnovane elemente digitalnih storitev, ki izkoriščajo kognitivne pristranskosti

²⁶ S. Benartzi, J. Lehrer (2015), str. 58–59.

²⁷ A. Kramer *et al.* (2014); V. Bakir (2020).

²⁸ J. Trzaskowski (2024); M. R. Leiser, M. Caruana (2021).

in psihološke ranljivosti z namenom, da uporabnika zavedejo v vedenja, ki jih ob ustrezni informiranosti ali večji refleksiji ne bi izbral. To, kar temne vzorce dela temne, je dejstvo, da so oblikovani na način, ki izkorišča kognitivne pristranskosti in psihološke ranljivosti uporabnikov, in jih tako poskušajo zmanipulirati, da ravnajo v nasprotju s svojimi interesi oziroma drugače, kot bi ravnali, če bi bili ustrezno informirani.

Temni vzorci, ki jih vse bolj množično uporabljajo spletni trgovci, družbena omrežja, industrija videoiger in aplikacije za mobilne naprave, spodbujajo uporabnike k razkrivanju osebnih podatkov, ohranjanju pozornosti na platformi ter celo k razvoju kompulzivnih vedenj. Njihova učinkovitost je empirično potrjena: eksperiment Lugini in Strahilevitz (2021) je pokazal, da so blagi temni vzorci več kot podvojili odstotek uporabnikov, ki so se naročili na dvomisljivo storitev, agresivni pa so ga povečali za kar štirikrat.²⁹

5. Normativna analiza: etični ugovori in pravni odzivi

Konceptualni okvir vedenjske ekonomije – zlasti teorija izbirne arhitekture in dregljajev – ni uporaben zgolj za razumevanje mehanizmov, prek katerih sodobne tehnologije vplivajo na naše vedenje, temveč tudi kot izhodišče za etično evalvacijo takšnega vplivanja. Teorija dregljajev in širše ideološko jedro libertarnega paternalizma sta sprožila živahne razprave o dopustnosti tovrstnih intervencij, zlasti v liberalno-demokratskih družbah, v katerih velja varstvo avtonomije posameznika za osrednjo vrednoto. Ugovore, ki so bili prvotno razviti v zvezi z analognimi dregljaji – očitke o kratenju avtonomije, omejevanju svobodne izbire, manipulaciji ter zasledovanju ciljev, ki niso v interesu uporabnika –, je mogoče uporabiti tudi pri normativni presoji digitalnih dregljajev.

5.1. Avtonomija

Največji skupni imenoalec očitkov zoper dregljaje je, da ti spodkopavajo avtonomijo, saj posamezniku jemljejo nadzor nad lastnim procesom odločanja in ga s subtilnimi ter netransparentnimi vplivi usmerjajo v izbire, ki ne izhajajo iz njegovih lastnih premislekov, vrednot in želja.³⁰ Mogoče je trditi, da se grožnja avtonomiji v kontekstu digitalnih dregljajev okrepi. Digitalno okolje namreč omogoča njihovo stalno prisotnost, kar zmanjšuje možnost »odmika« in refleksije, to pa dodatno omejuje sposobnost avtonomnega odločanja. Posebej problematični so personalizirani in adaptivni dregljaji, ki so zasnovani z namenom poseganja v procese odločanja posameznika in temeljijo na izraziti asimetriji moči. Digitalne platforme razpolagajo z velikimi bazami podatkov o uporabnikih in strokovnim znanjem za njihovo analizo, medtem ko posameznik pogosto nima dostopa niti do tega, kaj vse o njem zbirajo, kako se ti podatki obdelujejo in za katere vedenjske posege se nato uporabljajo. To ustvarja radikalno neenakost v odnosu moči, saj se uporabnik odloča v arhitekturi, ki jo v celoti nadzoruje nekdo drug.

Če digitalni dregljaji avtonomijo erodirajo predvsem s pogostimi in prilagojenimi pritiski na kognitivne bližnjice, temni vzorci pomenijo še globlji poseg: njihov cilj je sistematično izničiti možnost avtonomnega odločanja. Temni vzorci niso le tveganje za zmanjševanje avtonomije, ampak jo praviloma aktivno izigravajo, saj uporabnika usmerjajo k izbiram, ki

²⁹ J. Lugini, L. J. Strahilevitz (2021).

³⁰ L. Bovens (2009).

so v neposrednem nasprotju z njegovimi interesi. V tem smislu so temni vzorci skrajna točka kontinuuma vedenjskih intervencij – so primer intervencij, ki avtonomijo posameznika ne le ogrožajo, temveč jo dejansko instrumentalizirajo.

5.2. Svoboda izbire

Drugi klasičen ugovor zoper dregljaje je, da ogrožajo svobodo izbire. Že v analognem okolju so podporniki te teorije težko prepričali skeptike, da takšne intervencije ne zmanjšujejo nabora možnosti in da se jim je mogoče enostavno izogniti. V digitalnem okolju pa se ta očitek le še krepí. Digitalna arhitektura je zasnovana tako, da sistematično spodbuja intuitivno in avtomatično delovanje, zaradi česar je domnevna izogibljivost dregljajem pogosto zgolj teoretična. Čeprav digitalni dregljaji formalno ne omejujejo nabora možnosti, se teza o njihovi enostavni izogibljivosti v kontekstu okolja, ki spodbuja hitre in avtomatske odzive, izkaže za (bolj ali manj) naivno – zdi se celo, da se jim je mogoče v celoti izogniti le tako, da se izognemo digitalnemu okolju samemu.

Če digitalni dregljaji svobodo izbire omejujejo predvsem z oblikovanjem okolja, ki uporabnike potiska k določenim možnostim, pa grede temni vzorci še dlje. Ti so oblikovani prav z namenom, da izbiro izkrivijo ali prikrijejo, tako da uporabnik ne zazna dejanskega obsega možnosti ali pa ga dizajn namerno zavede k izbiri, ki mu škoduje. Medtem ko pri dregljajih ostaja vsaj formalna odprtost možnosti, temni vzorci to odprtost sistematično izničijo.

5.3. Manipulacija

S selitvijo iz analognega v digitalno okolje sta se okrepila tudi manipulativen potencial dregljajev ter možnost, da so ti uporabljeni za zasledovanje ciljev, ki niso v uporabnikovem, ampak »arhitektovem interesu«. Zasledovanje takšnih ciljev namreč pri digitalnih dregljajih ni le hipotetično, ampak je vgrajeno v samo naravo platformne digitalne ekonomije, ki temelji na barantanju s pozornostjo.³¹ Posebej spletna družbena omrežja, spletne igre itn. temeljijo na »ugrabitvi« pozornosti uporabnikov in povečevanju njihovega angažmaja na platformi. Brez uporabnikov in njihovih podatkov digitalne storitve ostanejo prazne »tržnice«.

Ponudniki spletnih storitev so že v izhodišču gnani k temu, da pridobijo, ohranjajo in povečujejo število uporabnikov, da njihove digitalne storitve ne (p)ostanejo okostja brez življenja. Možnost »zlobnega arhitekta« je zaradi človeško ustvarjene narave digitalnega »sveta« in ekonomskega okvira, v katerem digitalne storitve prihajajo na svet, zato v digitalnem okolju prisotna bolj kot v analognem. Digitalne storitve zato lahko razvrstimo na premici od manj do bolj manipulabilnih, same manipulabilnosti pa ne moremo izključiti. Močnejši in učinkovitejši kot je obvoz racionalnega premisleka in natančnejše kot je »igranje na (možganske) strune« uporabnikov, bolj je digitalna storitev manipulativna. Glede na to, da podjetja »živijo ali (pro)padejo« na številu uporabnikov, je neizbežno, da bodo znanje o uporabnikih tudi uporabila.

V primeru temnih vzorcev manipulacija ni več zgolj stranski učinek zasnove, temveč njen primarni namen. Temni vzorci so (po definiciji) oblikovani zato, da uporabnika zavedejo, zlorabijo njegove ranljivosti in ga usmerijo k dejanjem, ki so v nasprotju z njegovimi interesi. Če lahko digitalne dregljaje še vedno umestimo v sivo cono med dopustnim in problematičnim vedenjskim vplivanjem, temni vzorci to mejo presežejo in pomenijo nedvoumno manipulacijo.

³¹ T. Wu (2016).

5.4. (Ne)transparentnost

V kontekstu digitalnih dregljajev se posebej izostrti tudi ugovor netransparentnosti. Še zlasti pri personaliziranih in adaptivnih dregljajih problem ni zgolj v tem, da uporabniki ne prepoznajo vedenjskih mehanizmov, ki usmerjajo njihove odločitve, temveč tudi v dejstvu, da ostajajo prikriti načini zbiranja in obdelave osebnih podatkov, na podlagi katerih ti mehanizmi delujejo. Posameznik se tako ne odloča le v arhitekturi, ki je netransparentno zasnovana, temveč tudi na podlagi podatkovnih procesov, ki so mu v celoti nevidni.

6. Pravni odziv

Ker imajo številne odločitve posameznikov tudi pravne posledice, pravo določa (vsaj nekatere) meje dopustnega vplivanja in tako posameznika varuje tudi kot odločevalca. V civilnem pravu je veljavnost pogodbe vezana na poslovno sposobnost pogodbenih strank ter odsotnost zmot, grožnje, sile ali zvižaje; potrošniško pravo pa prepoveduje t. i. zavajajoče in agresivne prakse, saj bi te utegnile potrošnika napeljati k odločitvi, ki je sicer ne bi sprejel. Najbolj eksplicitno ščiti svobodo odločanja kazensko pravo, ki z inkriminacijo prisiljenja varuje posameznika pred tem, da bi ga kdo s silo ali grožnjo prisilil v določeno ravnanje ali opustitev. To so le nekateri primeri, kako pravni red v različnih sferah varuje našo avtonomijo odločanja. V tej razpravi pa me zanima, kako – in če sploh – je posameznik kot odločevalec pravno varovan v digitalnem okolju: ali so opisani mehanizmi vedenjskega vplivanja pravno dopustni ali prepovedani? Sledi zelo kratek pregled veljavne pravne ureditve obravnavanega področja.

Varovanje uporabnikov spletnih storitev pred manipulativnimi praksami se je v EU sprva razvijalo predvsem v dveh smereh – v okviru prava varstva osebnih podatkov ter prava varstva potrošnikov.³² Uredba EU 2016/679 Evropskega parlamenta in Sveta z dne 27. aprila 2016 o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov ter o razveljavitvi Direktive 95/46/ES (Splošna uredba o varstvu podatkov, GDPR)³³ temnih vzorcev sicer ne ureja neposredno, a ker ti uporabnike na netransparenten in nepošten način usmerjajo k privolitvi v obdelavo osebnih podatkov ter pri tem izkoriščajo njihove psihološke ranljivosti, je mogoče ugotoviti, da kršijo temeljna načela varstva podatkov, zlasti načelo poštenosti. Takšno razlago je nedavno potrdil tudi Evropski odbor za varstvo podatkov v zavezujoči odločitvi 2/2023, sprejeti pozneje v končni odločbi irskega nadzornega organa zoper TikTok Technologies.

Pravo varstva potrošnikov se srečuje z izzivom, da vedenjske intervencije, opisane v tem besedilu, pogosto delujejo v sivi coni med dopustnimi tehnikami trženjskega prepričevanja in očitno nedopustnimi oblikami vplivanja, kot sta prisila in zavajanje. Čeprav Direktiva 2005/29/ES Evropskega parlamenta in Sveta z dne 11. maja 2005 o nepoštenih poslovnih praksah podjetij v razmerju do potrošnikov na notranjem trgu (Direktiva o nepoštenih poslovnih praksah)³⁴ temnih vzorcev ne omenja izrecno, Evropska komisija zagovarja stališče, da je mogoče na podlagi njenih splošnih določb, zlasti prepovedi zavajajočih praks (6. in 7. člen), prepovedati tudi temne vzorce.

³² M. Brennncke (2023); M. Leiser (2022).

³³ UL L 119, 4. 5. 2016, str. 1–88.

³⁴ UL L 149, 11. 6. 2005, str. 22–39.

Prvi akt EU, ki izrecno omenja in definira temne vzorce, je Uredba (EU) 2022/2065 Evropskega parlamenta in Sveta z dne 19. oktobra 2022 o enotnem trgu digitalnih storitev in spremembi Direktive 2000/31/ES (Akt o digitalnih storitvah, DSA).³⁵ V uvodni določbi 67 jih opredeli, v 25. členu pa prepoveduje zasnovo vmesnikov, ki zavajajo ali manipulirajo uporabnike oziroma bistveno zmanjšujejo njihovo sposobnost sprejemati svobodne in informirane odločitve. Akt o digitalnih storitvah tako neposredno povezuje prepoved temnih vzorcev z varovanjem avtonomije uporabnikov. Najnovejši predpis na tem področju je Uredba (EU) 2024/1689 Evropskega parlamenta in Sveta z dne 13. junija 2024 o določitvi harmoniziranih pravil o umetni inteligenci in spremembi nekaterih zakonodajnih aktov Unije (Akt o umetni inteligenci, AI Act),³⁶ ki med prepovedanimi praksami na področju uporabe sistemov umetne inteligence izrecno navaja tudi subliminalne tehnike, torej vplive, ki delujejo onkraj naše zavesti. V točki a prvega odstavka 5. člena je prepovedana uporaba sistemov, ki nezavedno spreminjajo vedenje posameznika tako, da bistveno spodbližajo njegovo sposobnost informirane odločitve. Neodgovorjeno pa ostaja vprašanje, kako široko bo ta določba uporabljana v praksi in katere konkretne oblike vedenjskega vplivanja bo zajela kot prepovedane.

7. Sklep

Analiza je pokazala, da digitalni dregljaji in temni vzorci predstavljajo dve različni ravni vedenjskega vplivanja v digitalnem okolju. Kot subtilne oblike vedenjskega vplivanja digitalni dregljaji – vsaj na konceptualni ravni – ne zmanjšujejo števila možnosti, temveč preoblikujejo kontekst odločanja, zaradi česar ostajajo zunaj jasnega regulativnega okvira. Zaradi svoje formalne izogibljivosti in navidezne benignosti se praviloma ne štejejo za *apriori* zavajajoče ali prisilne, zato jih pravo (še) ne prepoveduje. Digitalni dregljaji tako za zdaj vztrajajo v sivi coni med dopustnim vplivanjem in prepovedano manipulacijo.

Kot primeren kazalnik za razlikovanje med (potencialno legitimno) prepričevalno tehnologijo in manipulativno tehnologijo se kažejo temni vzorci. To pojasni, zakaj so temni vzorci že deležni pravnega odziva, digitalni dregljaji pa še ne. Tisto, kar je pri dregljajih – analognih in digitalnih – zgolj hipotetična nevarnost, je pri temnih vzorcih povsem eksplicitno: ti nikoli niso namenjeni večanju blaginje uporabnika, temveč izključno interesom ponudnikov. Njihova invazivnost in prikritost presegata temeljne kriterije za upravičeno vedenjsko intervencijo – zlasti načela preglednosti, reverzibilnosti in spoštovanja avtonomije. V tem smislu temni vzorci niso zgolj posebna oblika digitalnega dreganja, temveč kvalitativno drugačen fenomen. Zato ni presenetljivo, da jih pravo EU že eksplicitno obravnava – od Splošne uredbe in Direktive o nepoštenih poslovnih praksah do novejših instrumentov, kot sta Akt o digitalnih storitvah in Akt o umetni inteligenci.

Digitalni dregljaji tako za zdaj ostajajo v sivi coni med dopustnim prepričevanjem in prepovedano manipulacijo, temni vzorci se kažejo kot jasen primer prestopa v sfero pravno nedopustnega. Gre za pojav, ki presega tehnični okvir oblikovanja uporabniških vmesnikov in zahteva nadaljnjo pravno-filozofsko refleksijo ter krepitev obstoječih regulativnih mehanizmov, saj trenutni okviri še niso zadostni za učinkovito omejevanje njihove vsakdanje prisotnosti.

³⁵ UL L 277, 27. 10. 2022, str. 1–102.

³⁶ UL L 2024/1689, 12. 7. 2024.

Literatura

- Ariely, Dan (2010). *Predvidljivo nerazumni*. Ljubljana: Mladinska knjiga.
- Bakir, Vian (2020). Psychological Operations in Digital Political Campaigns: Assessing Cambridge Analytica's Psychographic Profiling and Targeting. *Frontiers in Communication*, št. 5 (67), doi: 10.3389/fcomm.2020.00067.
- Benartzi, Shlomo, Lehrer, Jonah (2017). *The Smarter Screen: Surprising Ways to Influence and Improve Online Behavior*. London: Portfolio.
- Bovens, Luc (2009). *The Ethics of Nudge*. V: T. Grüne-Yanoff, S. O. Hansson (ur.), *Preference Change: Approaches from Philosophy, Economics and Psychology*. Berlin, Heidelberg: Springer Science & Business Media.
- Brenncke, Martin (2023). Regulating Dark Patterns. *Notre Dame Journal of International & Comparative Law*, št. 14(1), str. 39–79.
- Briggs, A. Rachael (2023). Normative Theories of Rational Choice: Expected Utility. V: E. N. Zalta, U. Nodelman (ur.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), <https://plato.stanford.edu/archives/win2023/entries/rationality-normative-utility/> (15. 5. 2025).
- Carr, Nicholas (2011). *Plitvine: kako internet spreminja naš način razmišljanja, branja in pomnjenja*. Ljubljana: Cankarjeva založba.
- Damasio, R. Antonio (2009). *Iskanje Spinoze - Veselje, žalost in čuteči možgani*. Ljubljana: Založba Krtina.
- Grüne-Yanoff, Till (2012). Old Wine in New Casks: Libertarian Paternalism Still Violates Liberal Principles. *Social Choice & Welfare*, št. 38(4), str. 635–45.
- Hausman, M. Daniel, Welch, Brynn (2010). Debate: To Nudge or Not to Nudge. *Journal of Political Philosophy*, št. 18(1), str. 123–36.
- Hummel, Dennis, Peyman Toreini, Maedche, Alexander (2018). *Improving Digital Nudging Using Attentive User Interfaces: Theory Development and Experiment Design*. Paper presented at DESRIST conference, Chennai, India.
- Kramer, Adam, Guillory, E. Jamie, Hancock, T. Jeffrey (2014). Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks. *Proceedings of the National Academy of Sciences*, št. 111(24), str. 8788–8790.
- Leiser, Mark, Caruana, Mireille (2021). Light to Be Found in Europe's Consumer Protection Regime. *Journal of European Consumer and Market Law*, št. 10(6), str. 237–51.
- Leiser, Mark (2022). *Dark Patterns: The Case for Regulatory Pluralism*. V: I. Kamara, R. Leenes, E. Kosta (ur.), *Research Handbook on EU Data Protection Law*. London: Taylor and Francis.
- Lembcke, Tim-Benjamin (2019). *Towards a Unified Understanding of Digital Nudging by Addressing Its Analog Roots*. Paper presented at Pacific Asia Conference on Information Society, XI'AN, China.
- Luguri, Jamie, Strahilevitz, Lior Jacob (2021). Shining a Light on Dark Patterns. *Journal of Legal Analysis*, št. 13(1), str. 43–109.
- Meske, Christian, Potthoff, Tobias (2017). *The dinu-model – a process model for the design of nudges*. Paper presented at European Conference on Information Systems (ECIS), Guimarães, Portugal.

- Rebonato, Riccardo (2012). *Taking Liberties: A Critical Examination of Libertarian Paternalism*. New York: Palgrave Macmillan.
- Schmidt, Andreas, Engelen, Bart (2020). The Ethics of Nudging: An Overview. *Philosophy Compass*, št. 15(4), str. e12658.
- Simon, Herbert A. (1957). *Models of man: Social and rational; mathematical essays on rational human behavior in a social setting*. New York: John Wiley & Sons.
- Slovic, Paul (1987). Perception of Risk. *Science*, št. 236(4799), str. 280–85.
- Susser, Daniel, Roessler, Beate, Nissenbaum, Helen. (2019). Technology, Autonomy, and Manipulation. *Internet Policy Review - Journal on Internet Regulation*, št. 8(2).
- Strle, Toma, Markič, Olga (2021). *O odločanju in osebni avtonomiji*. Maribor: Založba Aristej.
- Thaler, Richard H., Sunstein, Cass R. (2003a). Libertarian Paternalism. *The American Economic Review*, št. 93(2), str. 175–79.
- Thaler, Richard H., Sunstein, Cass R. (2003b). Libertarian Paternalism Is Not an Oxymoron. *The University of Chicago Law Review*, št. 70(4), str. 1159–202.
- Thaler, Richard H., Sunstein, Cass R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York: Penguin Books.
- Thaler, Richard H., Sunstein, Cass R. (2022). *Dregljaj – Izboljšanje odločitev o zdravju, blaginji in sreči*, dokončna izdaja. Ljubljana: UMco.
- Trzaskowski, Jan (2024). Manipulation by Design. *Electronic Markets*, št. 34(14).
- Tversky, Amos, Kahneman, Daniel (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, št. 185(4157), str. 1124–31.
- Tversky, Amos, Kahneman, Daniel (1981). The Framing of Decisions and the Psychology of Choice. *Science*, št. 211(4481), str. 453–58.
- Weinmann, Markus, Schneider, Christoph, vom Brocke, Jan (2016). Digital Nudging. *Business & Information Systems Engineering*, št. 58(6), str. 433–36.
- Wu, Tim (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. New York: Knopf.