

# Filozofski savjetnik

ČASOPIS ZA FILOZOFSKU PRAKSU · HRVATSKO UDRUŽENJE ZA FILOZOFSKU PRAKSU

SVEZAK Vol. I BROJ Br. I GODINA 2026 NAKLADNIK HUFPP, Zagreb

FILOZOFIJA · BIOETIKA · IZVORNI ZNANSTVENI ČLANAK

## Umjetna inteligencija u mentalnom zdravlju:

*Bioetički pogled i regulativa*

**Ivica Kelam** · *Fakultet za dentalnu medicinu i zdravlje, Sveučilište J. J. Strossmayera u Osijeku, Osijek*

**Filip Has** · *Fakultet za odgojne i obrazovne znanosti, Sveučilište J. J. Strossmayera u Osijeku, Osijek*

Primljeno: 26. 4. 2026 · Prihvaćeno: 11. 5. 2026

### SAŽETAK

Razvoj i primjena umjetne inteligencije sve više prodiru u različita područja djelovanja, uključujući i mentalno zdravlje, pri čemu se otvaraju brojne stručne, pravne i bioetičke dileme. U tom kontekstu osobito dolaze do izražaja regulatorni izazovi koji zahtijevaju osobitu pozornost s interdisciplinarnim pristupom. Među njima se ističu pitanja poput zaštite osobnih podataka, odgovornosti u slučaju pogrešnih ili štetnih psihoterapijskih preporuka te mogućnost zamjene interpersonalne interakcije automatiziranim sustavima u ovakvim osjetljivim odnosima. Nužna je uspostava jasnog i koherentnog regulatornog okvira uz uvažavanje bioetičkih vrijednosti. Primjenom kvalitativne analize relevantne literature i pregledom pozitivnopravnih propisa Republike Hrvatske, europskog Akta o umjetnoj inteligenciji, identificirati će se regulatorni izazovi te otvoriti prostor za normativno i bioetičko promišljanje o primjeni umjetne inteligencije u mentalnom zdravlju.

**KLJUČNE RIJEČI:** bioetika, psihoterapija, regulativa, Republika Hrvatska, umjetna inteligencija



## I. *Uvod*

---

Suvremeno društvo obilježeno je paradoksom koji se na prvi pogled teško uočava, ali se duboko osjeća u svakodnevnom životu: nikada nismo bili povezaniji, a istodobno nikada usamljeniji. Gradovi su sve gušći, komunikacija brža, mreže šire, ali između ljudi kao da se širi nevidljivi prostor tišine. U tom prostoru, pojedinac sve češće ostaje sam sa sobom. Ne nužno fizički izoliran, već emocionalno udaljen od drugih. Rad, zabava, komunikacija i informiranje - sve se odvija kroz zaslon. Ljudi više ne izlaze kako bi susreli druge; oni se povezuju bez dodira, bez pogleda, bez tišine koja se dijeli u istoj prostoriji. Usamljenost u takvom kontekstu ne izgleda kao praznina, već kao ispunjenost notifikacijama, porukama, sadržajem. Digitalni alati nude brzinu, dostupnost i osjećaj kontrole, ali ne mogu nadomjestiti složenost ljudskog odnosa. U stvarnom susretu postoji nepredvidivost, ranjivost i prisutnost drugoga; u digitalnom okruženju sve je posredovano, filtrirano i pojednostavljeno. Upravo u tom prostoru između potrebe za povezanošću i stvarne udaljenosti nastaje nova dinamika traženja pomoći. Kada se pojave emocionalne poteškoće, tjeskoba ili osjećaj izgubljenosti, pojedinac sve rjeđe spontano poseže za izravnim ljudskim kontaktom. Razlozi su različiti: stigma, nedostupnost stručnjaka, vremenska ograničenja ili jednostavno navika da se odgovori traže ondje gdje se i inače traži sve ostalo - na internetu. Digitalni alati tada ne predstavljaju samo pomoćno sredstvo, nego postaju prvi sugovornik. U takvom okruženju nije iznenađujuće što se ljudi okreću aplikacijama, chatbotovima i drugim oblicima digitalne podrške. Oni su uvijek dostupni i ne zahtijevaju izlaganje koje često prati susret s drugom osobom. U svijetu u kojem je sve ubrzano, digitalna pomoć nudi trenutni odgovor, makar i pojednostavljen. Ona pruža iluziju razumijevanja, kontinuiteta i sigurnosti. No iza te dostupnosti krije se i pitanje koje ostaje otvoreno: može li tehnologija uistinu zamijeniti odnos? Može li algoritam prepoznati nijanse tišine između riječi, ili težinu pogleda koji se ne vidi kroz ekran? Možda odgovor nije u potpunoj zamjeni, već u razumijevanju zašto ljudi sve češće biraju takav put. U društvu koje se polako zatvara unutar vlastitih prostora, digitalni oblici pomoći ne predstavljaju odstupanje, nego logičan nastavak načina života. U konačnici, ako su naši razgovori preseljeni na ekrane, nije neobično da su tamo preseljene i naše brige. Iako primjena umjetne inteligencije u mentalnom zdravlju otvara neke nove i korisne mogućnosti u praksi, otvara istodobno i niz složenih izazova koji nadilaze isključivo tehničku dimenziju, i zahtijevaju interdisciplinarno promišljanje na sjecištu prava, bioetike i psihoterapijske prakse.

### 1.1. Metodologija istraživanja

Ovaj rad temelji se na kvalitativnom i interdisciplinarnom istraživačkom pristupu koji obuhvaća analizu i sintezu relevantne literature te normativnu analizu pozitivnopravnih izvora, konkretnije europske Uredbe o umjetnoj inteligenciji i relevantnih propisa Republike Hrvatske. Dodatno, u radu su izdvojeni konkretni slučajevi iz prakse radi ilustracije bioetičkih izazova primjene umjetne inteligencije u mentalnom zdravlju. Kritičkom osvrtom na pojavne oblike primjene umjetne inteligencije u području mentalnog zdravlja i normativnom evaluacijom pozitivnopravnih izvora pruža se znanstveni doprinos s ciljem identifikacije bioetičkih i regulatornih izazova u Republici Hrvatskoje (i šire) te se otvara prostor za daljnja teorijska i praktična promišljanja.

## II. Umjetna inteligencija u psihoterapiji – Pojam, oblici i primjene

Rapidni razvoj umjetne inteligencije rezultirao je njezinom sve većom prisutnošću općenito u području zdravstvene skrbi.<sup>1</sup> Ona u najširem smislu označava skup tehnologija koje omogućuju računalnim sustavima izvršavanje zadataka koji inače zahtijevaju ljudsku kognitivnu sposobnost, poput učenja, zaključivanja, prepoznavanja obrazaca i donošenja odluka.<sup>2</sup> U kontekstu mentalnog zdravlja, umjetna inteligencija obuhvaća različite oblike algoritamskih sustava koji analiziraju podatke, predviđaju, generiraju preporuke ili sudjeluju u komunikaciji s korisnicima zdravstvenih usluga.<sup>3</sup> Kada je riječ o psihoterapiji, primjena umjetne inteligencije ne predstavlja jedinstven i homogen fenomen, već obuhvaća različite oblike tehnoloških rješenja koji se razlikuju prema svojoj funkciji, stupnju autonomije i ulozi u psihoterapijskom procesu.

Jednu skupinu čine sustavi koji omogućuju izravnu komunikaciju s korisnicima, najčešće u obliku tzv. psihoterapijskih „chatbotova“ ili virtualnih konverzijskih agenata. Njihova osnovna funkcija jest omogućiti interakciju koja u određenoj mjeri imitira ljudski razgovor, pri čemu se oslanjaju na kombinaciju tehnika obrade prirodnog jezika, strojnog učenja i unaprijed definiranih

<sup>1</sup> Aleksandra Krajnović, Monika Hordov, Zvonko Pešić, "Integrirana briga o zdravlju i holistički pristup - nova paradigma upravljanja zdravljem populacije uz primjenu nove tehnologije i umjetne inteligencije," u: *10. međunarodna znanstvena konferencija CRODMA* (Varaždin: Sveučilište u Zagrebu Fakultet organizacije i informatike, 2025), 201. Usp. Anamaria Malešević, Anto Čartolovni, "Umjetna inteligencija i donošenje odluka u medicini: pogled ključnih dionika zdravstvenog sustava," *Glasnik Hrvatskog katoličkog liječničkog društva* 34, br. 4 (2024): 296. Usp. Eric J. Topol, "High-performance medicine: the convergence of human and artificial intelligence", *Nature Medicine* 25, br. 1 (2019): 44.

<sup>2</sup> Stuart Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach* (Englewood Cliffs, New Jersey: Prentice Hall, 1995), 5.

<sup>3</sup> Simon D'Alfonso, "AI in mental health", *Current Opinion in Psychology* 36 (2020): 112-113. Usp. Eric J. Topol, *op. cit. u bilj.* 1, 47. Usp. Thomas Davenport i Ravi Kalakota, "The potential for artificial intelligence in healthcare", *Future Healthcare Journal* 6, br. 2 (2019): 94-96.

komunikacijskih modela. Takvi sustavi mogu pružati psihološku podršku, voditi strukturirane razgovore temeljene na određenim psihoterapijskim pristupima, pomagati u praćenju emocionalnog stanja korisnika i savjetovati.<sup>4</sup> Drugu skupinu čine sustavi za podršku odlučivanju i radu psihoterapeuta. Riječ je o alatima koji mogu analizirati podatke o korisnicima, prepoznati psihičke poteškoće, sugerirati moguće psihoterapijske pristupe, pomagati u procjeni rizika (primjerice suicidalnosti ili nasilnog ponašanja) ili strukturirati tijek psihoterapijskog procesa. U takvim slučajevima umjetna inteligencija ne djeluje kao zamjena za psihoterapeuta, već kao pomoćni alat, pri čemu odgovornost za donošenje konačnih odluka i dalje ostaje na stručnjaku.<sup>5</sup>

Neovisno o konkretnoj kategoriji, zajedničko obilježje navedenih oblika primjene umjetne inteligencije u mentalnom zdravlju jest njihovo djelovanje u području koje je obilježeno visokom razinom osjetljivosti, kako u pogledu prirode podataka koji se obrađuju, tako i u pogledu odnosa između psihoterapeuta i korisnika. Upravo ta specifičnost razlikuje psihoterapiju od mnogih drugih područja primjene umjetne inteligencije u zdravlju te zahtijeva dodatnu razinu pravne i bioetičke refleksije. Važno je razlikovati uporabu umjetne inteligencije kao već spomenutog pomoćnog alata u radu stručnjaka od njezine potencijalne uporabe kao zamjene za određene (ili sve) elemente psihoterapijskog procesa. Ta razlika ima neposredne implikacije ne samo na razinu prihvatljivosti takvih tehnologija u praksi, nego i na pitanja odgovornosti, autonomije korisnika i očuvanja temeljnih obilježja psihoterapijskog odnosa.

### III. *Specifičnost psihoterapijskog odnosa*

Psihoterapijski odnos predstavlja temeljni element psihoterapijskog procesa te se prepoznaje kao jedan od elementarnih čimbenika psihoterapijskog ishoda. Za razliku od mnogih drugih oblika zdravstvene skrbi, u kojima je naglasak primarno na dijagnostici i intervenciji, psihoterapija se u velikoj

---

<sup>4</sup> Marko Ćurković, "Primjena chatbotova u širem okviru skrbi za mentalno zdravlje", u: *Praktična primjena chatbotova u svakodnevnom (psihijatrijskom) radu*, Ur. Goran Arbanas (Zagreb: Klinika za psihijatriju Vrapče, Sveučilište u Zagrebu Medicinski fakultet, 2024), 34-36. Usp. Alaa Ali Abd-Alrazaq *et al.*, "Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis", *Journal of Medical Internet Research* 22, br. 7 (2019): 3. Usp. Becky Inkster, Shubhankar Sarda, Vinod Subramanian, "An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study", *JMIR mHealth and uHealth* 6, br. 11 (2018): 2-3. Usp. Kathleen K. Fitzpatrick, Alison Darcy, Molly Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial", *JMIR Mental Health* 4, br. 2 (2017): 2-3.

<sup>5</sup> Sarah Graham *et al.*, "Artificial intelligence for mental health and mental illnesses: an overview", *Current Psychiatry Reports* 21, br. 11 (2019): 3-5. Usp. Adrian B. R. Shatte, Delyse M. Hutchinson, Samantha J. Teague, "Machine learning in mental health: a scoping review of methods and applications", *Psychol Med.* 49, br. 9 (2019): 1426-1448; David D. Luxton, "Artificial intelligence in psychological practice: current and future applications and implications", *Professional Psychology: Research and Practice* 45, br. 5 (2014): 3.

mjeri oslanja na kvalitetu interpersonalne interakcije između psihoterapeuta i osobe uključene u psihoterapijski proces.<sup>6</sup>

Središnji pojam u tom kontekstu jest tzv. „psihoterapijski savez“, koji obuhvaća uzajamno povjerenje, empatiju, dogovor o ciljevima psihoterapije i suradnju u ostvarivanju tih ciljeva. Istraživanja pokazuju da kvaliteta takvog odnosa predstavlja jedan od najsnažnijih prediktora uspješnosti psihoterapije, neovisno o konkretnom psihoterapijskom pravcu.<sup>7</sup> Upravo zbog toga, svaka intervencija koja potencijalno utječe na prirodu i sadržaj tog odnosa zahtijeva posebnu pažnju. Psihoterapijski odnos obilježen je visokom razinom povjerljivosti te obvezom čuvanja profesionalne tajne, što je istodobno pravni i etički standard. Povjerljivost ne predstavlja samo formalnu obvezu, već i preduvjet za uspostavu sigurnog prostora u kojem osoba može slobodno iznositi osobna iskustva, emocionalne poteškoće i osjetljive životne okolnosti. U tom smislu, povjerenje u psihoterapeuta i sigurnost komunikacijskog okruženja imaju izravne implikacije na učinkovitost psihoterapijskog procesa. Postoji stoga rizik da pretjerano oslanjanje na tehnološka rješenja dovede do postupnog potiskivanja interpersonalne dimenzije psihoterapije, čime se narušava njezina temeljna esencija. Iz bioetičke perspektive, takav razvoj događaja zahtijeva oprez, jer očuvanje ljudskog dostojanstva i integriteta osobe predstavlja središnje načelo svake zdravstvene skrbi, a osobito psihoterapije. Dodatna specifičnost psihoterapijskog odnosa proizlazi iz činjenice da se često radi o osobama koje se nalaze u stanju povećane ranjivosti, bilo zbog psihičkih poteškoća, emocionalne nestabilnosti ili životnih kriza.<sup>8</sup> Takva ranjivost povlači za sobom element asimetrije u odnos između psihoterapeuta i osobe u psihoterapiji, pri čemu psihoterapeut raspoložuje stručnim znanjem i profesionalnim autoritetom, dok druga strana u određenoj mjeri ovisi o njegovoj procjeni i vođenju psihoterapijskog procesa.

Upravo navedene karakteristike - psihoterapijski savez, povjerljivost, ranjivost i asimetrija odnosa - čine psihoterapiju osobito osjetljivim područjem za povezivanje s umjetnom inteligencijom. Za razliku od tehnoloških rješenja u drugim područjima zdravstvene skrbi, gdje se umjetna inteligencija često koristi kao alat za analizu podataka ili podršku dijagnostici, u psihoterapiji

---

<sup>6</sup> John C. Norcross, Michael J. Lambert, "Psychotherapy relationships that work III", *Psychotherapy* 55, br. 4 (2018): 313. Usp. Edward S. Bordin, "The generalizability of the psychoanalytic concept of the working alliance", *Psychotherapy: Theory, Research & Practice* 16, br. 3 (1979): 253.

<sup>7</sup> Christoph Flückiger *et al.*, "The alliance in adult psychotherapy: a meta-analytic synthesis", *Psychotherapy* 55, br. 4 (2018): 317-319. Usp. Adam O. Horvath *et al.*, "Alliance in individual psychotherapy", *Psychotherapy* 48, br. 1 (2011): 10-11. Usp. Edward S. Bordin, *op. cit. u bilj.* 6, 252.

<sup>8</sup> Tom L. Beauchamp, James F. Childress, *Principles of Biomedical Ethics*, 8. izd. (Oxford: Oxford University Press, 2019), 89-90.

ona potencijalno zadire u samu srž interpersonalnog odnosa. Time se otvara pitanje može li tehnološki sustav, bez obzira na stupanj svoje sofisticiranosti, adekvatno replicirati ili nadomjestiti elemente poput empatije, razumijevanja i autentične ljudske interakcije.<sup>9</sup> Osim toga, uvođenje umjetne inteligencije u psihoterapijski odnos može utjecati i na percepciju odgovornosti, kako s aspekta stručnjaka tako i s aspekta osobe uključene u psihoterapijski proces. Ako se određeni elementi psihoterapije delegiraju algoritamskim sustavima, postavlja se pitanje u kojoj mjeri to mijenja tradicionalno shvaćanje profesionalne odgovornosti psihoterapeuta te može li doći do razvodnjavanja odgovornosti u slučaju štetnih posljedica. Slijedom navedenoga, psihoterapijski odnos ne može se promatrati kao neutralan teren za puku implementaciju tehnoloških rješenja, već kao kompleksan bioetički i profesionalni prostor u kojem svaka intervencija mora biti pažljivo evaluirana. S bioetičkog stajališta, načela neškodljivosti i odgovornosti zahtijevaju jasnu raspodjelu odgovornosti i uspostavu mehanizama nadzora koji će osigurati zaštitu korisnika.

#### IV. *Studije o spornim primjerima iz prakse*

Znanstvena literatura i dokumentirani slučajevi pokazuju da primjena umjetne inteligencije u području mentalnog zdravlja nije samo teorijsko pitanje, nego i područje u kojem već postoje prednosti ali i konkretni problemi. Sustavni pregled etičkih izazova umjetne inteligencije u mentalnom zdravlju identificirao je kao najčešće teme sigurnost i štetu, privatnost i povjerljivost, odgovornost, transparentnost i povjerenje te rizik od ovisnosti o takvim sustavima.<sup>10</sup> Rahsepar Meadi i sur. navode primjer u kojem je korisnik, predstavljajući se kao dijete, chatbotu prijavio da je bio spolno zlostavljan, a sustav je odgovorio na način koji nije prepoznao ozbiljnost situacije, nego je dao neprimjeren odgovor, što autori tumače kao posljedicu nemogućnosti chatbota da kontekstualizira korisničke znakove i upravlja krizom. U istom istraživanju istaknut je i slučaj wellness chatbota Tessa, koji je nakon davanja štetnih savjeta o mršavljenju korisnicima s poremećajima hranjenja uklonjen s platforme Nacionalne udruge za poremećaje prehrane u SAD-u, što je poslužilo kao snažan argument za potrebu izrade posebnih etičkih smjernica za konverzijsku umjetnu inteligenciju u mentalnom zdravlju. Qiu i sur. upućuju na tragičan događaj u kojem je maloljetna osoba, nakon interakcije s

<sup>9</sup> P. Joseph Akhil, Babu Anithamol, "The unseen dilemma of AI in mental healthcare", *AI & Society* 40, br. 3 (2025): 1533. Usp. Liliana Laranjo *et al.*, "Conversational agents in healthcare: a systematic review", *Journal of the American Medical Informatics Association* 25, br. 9 (2018): 1249. Usp. David D. Luxton, *op. cit. u bilj.* 5, 4.

<sup>10</sup> Mehran Rahsepar Meadi *et al.*, "Exploring the ethical challenges of conversational AI in mental health care: scoping review", *JMIR Mental Health* 12, br. 1 (2025): 4-9.

chatbotom, počinila samoubojstvo, pri čemu je sustav navodno poticao suicidalne misli i ponašanje.<sup>11</sup> Ovakvi primjeri jasno ilustriraju potencijalne granice i opasnosti primjene umjetne inteligencije u području mentalnog zdravlja, osobito u kontekstu izostanka adekvatnog nadzora i zaštitnih mehanizama.

Osim sadržaja odgovora, sporna je i sama dinamika odnosa koji korisnici mogu razviti s takvim sustavima. Vaidyam i sur. podsjećaju da je još u ranim fazama razvoja chatbotova zabilježeno da vrlo kratka izloženost relativno jednostavnom programu može potaknuti snažnu emocionalnu vezanost i deluzivno pripisivanje sposobnosti razumijevanja, pri čemu su se neki korisnici emocionalno vezali uz sustav ELIZU i doživljavali je pouzdanijom nego što jest.<sup>12</sup> Isti autori ističu da su preliminarni rezultati o psihijatrijskoj uporabi chatbotova povoljni, ali da i dalje postoje ozbiljna otvorena pitanja u pogledu povjerljivosti, privatnosti, sigurnosti, odgovornosti i licenciranja stručnjaka koji takve sustave nadziru. Problematicnim se pokazuje i širenje modela digitalnog fenotipiziranja, odnosno prikupljanja tzv. digitalnih otisaka korisnika putem pametnih telefona, nosivih uređaja i društvenih mreža radi procjene mentalnog stanja. D'Alfonso upozorava da se suvremeni pristupi umjetne inteligencije u mentalnom zdravlju u velikoj mjeri oslanjaju upravo na rudarenju takvih podataka, uključujući jezične obrasce i podatke iz svakodnevne digitalne interakcije, što otvara prostor za osjetljive oblike nadzora i zaključivanja o mentalnom zdravlju pojedinca.<sup>13</sup> U tom smislu, psihoterapijski proces uključuje razmjenu izrazito osjetljivih informacija o mentalnom zdravlju, osobnim iskustvima i životnim okolnostima pojedinca, što prema Vayeni i sur. otvara problematična pitanja vezano uz privatnost, nadzor, odgovornost i javnog povjerenja u takve sustave.<sup>14</sup> Analize digitalnih aplikacija predviđene za mentalno zdravlje ukazuju na to da takvi sustavi često prikupljaju velike količine osjetljivih podataka, uključujući informacije o emocionalnom stanju i osobnim iskustvima, pri čemu politike privatnosti nisu uvijek transparentne niti lako razumljive korisnicima.<sup>15</sup> Dodatne tehničke analize ukazuju na to da određene aplikacije čak i dijele podatke s trećim stranama i ne osiguravaju adekvatne sigurnosne

---

<sup>11</sup> Jiahao Qiu *et al.*, "EmoAgent: Assessing and Safeguarding Human-AI Interaction for Mental Health Safety", u: *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Suzhou: Association for Computational Linguistics, 2025), 11741-11742.

<sup>12</sup> Aditya N. Vaidyam *et al.*, "Chatbots and conversational agents in mental health: a review of the psychiatric landscape", *The Canadian Journal of Psychiatry* 64, br. 7 (2019): 457. i 463.

<sup>13</sup> Simon D'Alfonso, *op. cit. u bilj.* 3, 112-113.

<sup>14</sup> Effy Vayena *et al.*, "Digital health: meeting the ethical and policy challenges", *Swiss Medical Weekly* 148 (2018): 3-6.

<sup>15</sup> Romael Haque, Sabirat Rubya, "For an app supposed to make its users feel better, it sure is a joke - an analysis of user reviews of mobile mental health applications", u: *Proceedings of the ACM on Human-Computer Interaction* 6 (New York: Association for Computing Machinery, 2022): 13-14.

standarde<sup>16</sup>, što predstavlja ozbiljan problem u kontekstu povjerljivosti koja je jedan od temelja psihoterapijskog odnosa. Iz bioetičke perspektive, zaštita privatnosti izravno je povezana s poštovanjem autonomije osobe, budući da informirani pristanak pretpostavlja jasno razumijevanje načina na koji se podaci koriste i obrađuju.

Mnogi sustavi umjetne inteligencije predstavljali su se kao alati za mentalno zdravlje, iako po svojim karakteristikama ne bi zadovoljavali standarde profesionalne psihoterapijske prakse<sup>17</sup>, što je moglo dovesti do pogrešnih očekivanja korisnika i odgode traženja prijeko potrebne stručne pomoći. Prethodno navedeni slučajevi potvrđuju upozorenja iz literature da digitalni alati mogu biti korisni kao dopuna odnosno pomoćno sredstvo stručnjacima, ali ne i kao potpuna zamjena za stručnu pomoć.<sup>18</sup> Sve izdvojeno potvrđuje da sporne primjene umjetne inteligencije u mentalnom zdravlju nisu ograničene samo na pitanje tehničke učinkovitosti, nego zadiru u samu jezgru povjerljivosti, autonomije, sigurnosti i profesionalne odgovornosti. Zaključno, odgovarajuća primjena umjetne inteligencije u mentalnom zdravlju pretpostavlja potrebu za postojanjem jasnih pravila o njenoj primjeni u praksi.<sup>19</sup>

## V. *AI ACT*

Na razini Europske unije regulatorni okvir primjene umjetne inteligencije značajno je unaprijeđen donošenjem Uredbe (EU) od 13. lipnja 2024. o utvrđivanju usklađenih pravila o umjetnoj inteligenciji<sup>20</sup> (dalje: AI Act), koja je stupila na snagu 1. kolovoza 2024. Riječ je o prvom sveobuhvatnom pravnom aktu kojim se na razini Unije uređuju razvoj, stavljanje na tržište, stavljanje u upotrebu i korištenje sustava umjetne inteligencije, s ciljem istodobnog poticanja inovacija i osiguravanja visoke razine zaštite zdravlja, sigurnosti i temeljnih prava.<sup>21</sup> AI Act temelji se na pristupu

<sup>16</sup> Leonardo Horn Iwaya *et al.*, "On the privacy of mental health apps: an empirical investigation and its implications for app development", *Empirical Software Engineering* 28, br. 1 (2023): 27-33.

<sup>17</sup> Effy Vayena *et al.*, *op. cit. u bilj.* 14, 1. Usp. John Torous, Laura Weiss Roberts, "Needed innovation in digital health and smartphone applications for mental health: transparency and trust", *JAMA Psychiatry* 74, br. 5 (2017): 438.

<sup>18</sup> Francesca Minerva, Alberto Giubilini, "Is AI the future of mental healthcare?", *Topoi* 42, br. 3 (2023): 815.

<sup>19</sup> Zeshan Iftikhar *et al.*, "How LLM counselors violate ethical standards in mental health practice: a practitioner-informed framework", u: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 8, br. 2 (Madrid: Association for the Advancement of Artificial Intelligence, 2025), 1320. Usp. David B. Olawade *et al.*, "Enhancing mental health with artificial intelligence: current trends and future prospects", *Journal of Medicine, Surgery, and Public Health* 3 (2024): 6-7. Usp. Nikolaos Koutsouleris *et al.*, "From promise to practice: towards the realisation of AI-informed mental health care", *The Lancet Digital Health* 4, br. 11 (2022): 837.

<sup>20</sup> Uredba (EU) 2024/1689 Europskog parlamenta i Vijeća od 13. lipnja 2024. o utvrđivanju usklađenih pravila o umjetnoj inteligenciji i o izmjeni uredaba (EZ) br. 300/2008, (EU) br. 167/2013, (EU) br. 168/2013, (EU) 2018/858, (EU) 2018/1139 i (EU) 2019/2144 te direktiva 2014/90/EU, (EU) 2016/797 i (EU) 2020/1828 (Akt o umjetnoj inteligenciji) (Tekst značajan za EGP) (Službeni list Europske unije L2024/1689) (dalje: AI Act)

<sup>21</sup> AI Act, preambula t. 1., 2. i 8., čl. 1.

koji počiva na procjeni rizika, pri čemu razlikuje više kategorija sustava umjetne inteligencije ovisno o razini rizika koji predstavljaju za temeljna prava i sigurnost pojedinaca. U tom smislu posebno se izdvajaju zabranjeni sustavi, sustavi visokog rizika te sustavi za koje su predviđene određene obveze transparentnosti.<sup>22</sup> Pri tome valja imati na umu da se AI Act ne primjenjuje na sve digitalne ili algoritamske sustave, već samo na one koji ispunjavaju definiciju umjetne inteligencije iz čl. 3. st. 1. AI Acta-a, što dodatno potvrđuju i smjernice Europske komisije o definiciji sustava umjetne inteligencije koje naglašavaju potrebu funkcionalnog i kontekstualnog tumačenja tog pojma.<sup>23</sup> Prema toj definiciji, sustav umjetne inteligencije bi bio strojni sustav dizajniran za rad s promjenjivim razinama autonomije, koji može pokazati prilagodljivost nakon uvođenja te koji, za eksplicitne ili implicitne ciljeve, iz ulaznih vrijednosti zaključuje kako generirati izlazne vrijednosti poput predviđanja, sadržaja, preporuka ili odluka koje mogu utjecati na fizička ili virtualna okruženja. Komisija dodatno pojašnjava da ta definicija obuhvaća sedam glavnih elemenata te da nije nužno svaka digitalna ili algoritamska aplikacija automatski i sustav umjetne inteligencije u smislu AI Acta.<sup>24</sup> To je za ovu tematiku važno zato što se psihoterapijska praksa danas može služiti različitim digitalnim alatima, ali samo dio njih mogao bi ulaziti u regulatorni doseg AI Acta.

Zabranjene prakse uređene su čl. 5. AI Act-a te obuhvaćaju sustave koji se smatraju neprihvatljivima zbog svojeg ozbiljnog utjecaja na temeljna prava. U kontekstu psihoterapije osobito su relevantne odredbe čl. 5. st. 1. t. (a) i (b), kojima se zabranjuju sustavi umjetne inteligencije koji koriste tehnike djelovanja izvan svjesne percepcije korisnika ili namjerno manipulativne ili obmanjujuće tehnike, kao i sustavi koji iskorištavaju ranjivosti pojedinaca, ako takve prakse imaju za cilj ili učinak bitno iskrivljavanje ponašanja te mogu uzrokovati znatnu štetu. Iako AI Act ne uređuje izričito psihoterapiju kao zasebno područje, navedene odredbe mogu biti relevantne u kontekstu primjene umjetne inteligencije u mentalnom zdravlju, osobito imajući u vidu da se psihoterapijski proces često odvija u okolnostima povećane psihološke ranjivosti. Upravo se u toj točki otvara najosjetljivija veza između AI Acta i psihoterapije. Psihoterapija je po svojoj naravi usmjerena na promjenu mišljenja, emocija i ponašanja, ali se ta promjena u legitimnoj psihoterapijskoj praksi temelji na profesionalnoj odgovornosti, informiranosti, autonomiji i psihoterapijskom savezu. Smjernice Europske komisije o zabranjenim praksama u području umjetne inteligencije dodatno pojašnjavaju da je središnje obilježje zabranjenih praksi upravo „bitno iskrivljavanje ponašanja“, odnosno situacija u

---

<sup>22</sup> *Ibid.* čl. 5., čl. 6. i čl. 50.

<sup>23</sup> Smjernice Komisije o definiciji sustava umjetne inteligencije utvrđenoj u Uredbi (EU) 2024/1689 (C(2025) 5053), čl. 1. st. 2.

<sup>24</sup> *Ibid.*, čl. 3. st. 61.

kojoj se značajno narušava sposobnost osobe da donese informiranu i autonomnu odluku.<sup>25</sup> U tom smislu može se postaviti pitanje mogu li određeni oblici primjene umjetne inteligencije, osobito oni koji uključuju automatizirano savjetovanje ili usmjeravanje ponašanja, u pojedinim okolnostima doći u dodir s granicama koje postavlja čl. 5. AI Act-a. Međutim, takvo tumačenje treba biti pod oprezom, budući da AI Act ne zabranjuje uporabu umjetne inteligencije u psihoterapijskom kontekstu kao takvu, već samo one prakse koje ispunjavaju stroge uvjete manipulacije, iskorištavanja ranjivosti i uzrokovanja štete.

Uz zabranjene prakse, AI Act posebno uređuje i tzv. visokorizične sustave umjetne inteligencije. Takvi sustavi, koji mogu imati značajan utjecaj na prava i sigurnost pojedinaca, podliježu nizu zahtjeva, uključujući uspostavu sustava upravljanja rizicima, osiguranje kvalitete podataka, transparentnost, tehničku dokumentaciju, kibernetičku sigurnost te obvezu ljudskog nadzora.<sup>26</sup> Iako psihoterapija nije izričito navedena među područjima visokog rizika, određeni sustavi koji se koriste u mentalnom zdravlju, ovisno o svojoj funkciji i učinku, mogu potpasti pod taj režim.<sup>27</sup> U tom smislu, za potrebe pravne analize osobito je važno razlikovati umjetnu inteligenciju kao pomoćni alat u radu stručnjaka od umjetne inteligencije koja preuzima funkcije koje bitno utječu na psihoterapijski proces ili na ponašanje osobe uključene u njega.

AI Act nije samodostatan regulatorni okvir. AI Act ima horizontalni karakter te se primjenjuje ne dovodeći u pitanje druge grane prava Unije, osobito pravo zaštite osobnih podataka, zaštite korisnika i temeljnih prava.<sup>28</sup> To znači da se njegova primjena u području psihoterapije nužno mora promatrati u vezi s postojećim pravnim okvirom, uključujući pravila o povjerljivosti, zaštiti podataka i profesionalnoj odgovornosti. U području psihoterapije to podrazumijeva da se AI Act nužno mora čitati zajedno s Općom uredbom o zaštiti podataka (dalje: GDPR)<sup>29</sup> i nacionalnim pravilima o profesionalnoj tajni, informiranom pristanku, pravima pacijenata i obavljanju psihoterapijske djelatnosti. Drugim riječima, i kada određeni sustav ne bi bio zabranjen prema čl. 5. AI Act-a niti bi očito potpadao pod kategoriju visokorizičnih sustava, njegova uporaba u psihoterapijskom kontekstu

---

<sup>25</sup> Smjernice Komisije o zabranjenim praksama u području umjetne inteligencije uspostavljenima Uredbom (EU) 2024/1689 (C(2025) 5052), čl. 3.2. i 3.3.

<sup>26</sup> AI Act, čl. 8.-15.

<sup>27</sup> *Ibid.*, čl. 6. i Prilog III. t. 1. (c) i t. 5. (a).

<sup>28</sup> AI Act, premabula t. 9. i 10.

<sup>29</sup> Uredba (EU) 2016/679 Europskog parlamenta i Vijeća od 27. travnja 2016. o zaštiti pojedinaca u vezi s obradom osobnih podataka i o slobodnom kretanju takvih podataka te o stavljanju izvan snage Direktive 95/46/EZ (Opća uredba o zaštiti podataka) (Tekst značajan za EGP) (Službeni list Europske unije L119/1) (dalje: GDPR)

i dalje može biti pravno i etički problematična s obzirom na posebnu osjetljivost podataka, ranjivost korisnika i zahtjev očuvanja autonomije i povjerljivosti.

Slijedom svega navedenoga, može se zaključiti da AI Act predstavlja važan i nužan korak u uspostavi europskog regulatornog okvira za umjetnu inteligenciju. AI Act stupio je na snagu dvadesetoga dana od dana objave u Službenom listu Europske unije, dok se općenito počinje primjenjivati od 2. kolovoza 2026., pri čemu je predviđena fazna primjena pojedinih njezinih dijelova: poglavlja I. i II. primjenjuju se od 2. veljače 2025., dok se poglavlje III. odjeljak 4., poglavlja V., VII. i XII. te članak 78. primjenjuju od 2. kolovoza 2025., uz iznimku članka 101., a članak 6. stavak 1. i s njime povezane obveze od 2. kolovoza 2027.<sup>30</sup> Premda je riječ o uredbi koja je u cijelosti obvezujuća i izravno se primjenjuje u svim državama članicama, države članice nisu oslobođene normativnih i institucionalnih obveza, nego moraju osigurati nacionalni provedbeni okvir: svaka država članica dužna je uspostaviti ili odrediti barem jedno notifikacijsko tijelo i barem jedno tijelo za nadzor tržišta kao nacionalna nadležna tijela, osigurati njihovu neovisnost te im pružiti dostatne tehničke, financijske i ljudske resurse i odgovarajuću infrastrukturu<sup>31</sup>, Komisiji priopćiti identitet tih tijela i njihove zadaće te najkasnije do 2. kolovoza 2025. javno objaviti podatke za kontakt nadležnih tijela i jedinstvene kontaktne točke<sup>32</sup>, a do istoga datuma i svake dvije godine nakon toga izvijestiti Komisiju o stanju njihovih financijskih i ljudskih resursa uz ocjenu njihove adekvatnosti<sup>33</sup>. Nadalje, države članice moraju utvrditi pravila o sankcijama i drugim provedbenim mjerama za povrede AI Act-a te poduzeti sve potrebne mjere kako bi se one pravilno i učinkovito provodile, pri čemu su o tim pravilima dužne obavijestiti Komisiju najkasnije do dana početka primjene AI Act-a, odnosno do 2. kolovoza 2026.<sup>34</sup> Iz navedenoga proizlazi da AI Act ne zahtijeva transpoziciju u nacionalna prava, ali zahtijeva pravodobnu institucionalnu, organizacijsku i sankcijsku prilagodbu nacionalnih pravnih poredaka radi njegove pune i učinkovite provedbe.

## VI. *Pravni okvir Republike Hrvatske*

U hrvatskom pravnom poretku ne postoji jedinstven i specifičan normativni okvir koji bi izravno uređivao primjenu umjetne inteligencije u psihoterapiji. Međutim, određeni aspekti relevantni za ovu problematiku uređeni su kroz više propisa koji se odnose na djelatnost psihoterapije, zaštitu

---

<sup>30</sup> AI Act, čl. 113.

<sup>31</sup> *Ibid.*, čl. 70. st. 1. i 3.

<sup>32</sup> *Ibid.*, čl. 70. st. 2.

<sup>33</sup> *Ibid.*, čl. 70. st. 6.

<sup>34</sup> AI Act, čl. 99. st. 1. i 2.

prava pacijenata te zaštitu osobnih podataka. Psihoterapija kao djelatnost uređena je Zakonom o djelatnosti psihoterapije<sup>35</sup>, koji propisuje uvjete za obavljanje psihoterapijske djelatnosti, stručne standarde, obvezu čuvanja profesionalne tajne te odgovornost psihoterapeuta.

Važnost u kontekstu mentalnog zdravlja ima i Zakon o zaštiti osoba s duševnim smetnjama<sup>36</sup>, koji polazi od zaštite dostojanstva i temeljnih prava osoba s duševnim smetnjama te propisuje da se medicinski postupci moraju primjenjivati u najboljem interesu osobe i na način koji u najmanjoj mjeri ograničava njezina prava.<sup>36</sup> Navedene odredbe mogu biti osobito relevantne u kontekstu primjene sustava umjetne inteligencije u psihoterapiji, budući da prekomjerna automatizacija odnosno depersonalizacija psihoterapijskog procesa upućuje na mogućnost dovođenja u pitanje individualizirani pristup i zaštitu prava ranjivih osoba. Navedeni zakoni, međutim, ne sadrže odredbe koje bi se konkretno odnosile na uporabu digitalnih alata ili sustava umjetne inteligencije u psihoterapijskom procesu, niti razrađuju granice njihove dopuštene primjene.

Navedeni zakon, međutim, ne sadrži odredbe koje bi se odnosile na uporabu digitalnih alata ili sustava umjetne inteligencije u psihoterapijskom procesu, niti razrađuje granice njihove dopuštene primjene. Drugim riječima, možemo tvrditi kako ovakva regulativa štiti interpersonalni element procesa psihoterapije. Zaštita prava osoba uključenih u psihoterapijski proces dodatno je uređena Zakonom o zaštiti prava pacijenata<sup>37</sup>, koji osigurava pravo na informirani pristanak, povjerljivost podataka i dostojanstvo pacijenta. U kontekstu primjene umjetne inteligencije, osobito je relevantno pravo pacijenta na obaviještenost, koje uključuje pravo na razumljive informacije o prirodi i načinu pružanja psihoterapijske usluge, što potencijalno može obuhvatiti i informacije o uporabi algoritamskih sustava u psihoterapijskom procesu. Nadalje, obrada osobnih podataka u psihoterapiji podliježe pravilima GDPR-a, koja podatke o zdravlju svrstava među posebne kategorije osobnih podataka čija je obrada načelno zabranjena, osim ako su ispunjeni strogi uvjeti zakonitosti obrade.<sup>38</sup> Iz navedenoga proizlazi da hrvatski pravni okvir pruža određenu razinu zaštite kroz opće institute zdravstvenog i podatkovnog prava, ali ne nudi izravne odgovore na pitanja koja proizlaze iz potencijalne primjene umjetne inteligencije u psihoterapijskom odnosu. Slijedom donošenja AI Act-a, pred hrvatskim zakonodavstvom nedvojbeno predstoji potreba za uspostavom odgovarajućih institucionalnih i normativnih okvira njegove provedbe, kao i za postupnom prilagodbom postojećeg zakonodavstva.

<sup>35</sup> Zakon o djelatnosti psihoterapije (NN 18/22)

<sup>36</sup> Zakon o zaštiti osoba s duševnim smetnjama (NN 76/14)

<sup>37</sup> Zakon o zaštiti prava pacijenata (NN 169/04, 37/08)

<sup>38</sup> GDPR, čl. 9.

## VI. *Zaključak*

Ideja da bi psihoterapijski proces u cijelosti mogao biti prepušten automatiziranim sustavima ne samo da je trenutno teško zamisliva, nego predstavlja ozbiljnu opasnost od dehumanizacije psihoterapijskog odnosa i gubitka njegovih temeljnih obilježja. Alati umjetne inteligencije mogu imati opravdanu i korisnu ulogu kao pomoćni alati u psihoterapijskoj praksi, primjerice u analizi podataka, praćenju stanja korisnika ili podršci odlučivanju psihoterapeuta. Međutim, njihova primjena ne može se promatrati izvan konteksta ranjivosti osoba uključenih u psihoterapijski proces, niti izvan profesionalne odgovornosti onih koji su ovlašteni da provode psihoterapijske postupke. Upravo u tom prostoru dolazi do izražaja temeljna bioetička dilema: u kojoj mjeri tehnološki sustavi mogu sudjelovati u procesima koji neposredno utječu na mentalno zdravlje pojedinca, a da pritom ne naruše autonomiju, integritet, dostojanstvo i sigurnost osobe?

U normativnom smislu, donošenje AI Acta predstavlja važan korak u uspostavi zajedničkog europskog regulatornog okvira. Iako odredbe o zabranjenim praksama, osobito one koje se odnose na manipulaciju i iskorištavanje ranjivosti, mogu imati značajnu ulogu kao zaštitni mehanizam, AI Act sam po sebi ne daje cjelovit odgovor na sva pitanja koja proizlaze iz primjene umjetne inteligencije u psihoterapiji. U tom smislu, može se smatrati važnom normativnom polazišnom točkom, ali ne i dovršenim rješenjem. Pred državama članicama Europske unije, uključujući Republiku Hrvatsku, tek predstoji razdoblje prilagodbe zakonodavstva, institucionalnih mehanizama i profesionalnih standarda novim tehnološkim okolnostima. Taj proces ne bi se trebao svesti isključivo samo na formalno usklađivanje s normama, već zahtijeva dublje promišljanje granica dopuštene primjene umjetne inteligencije u osjetljivim područjima poput mentalnog zdravlja. Polazeći od temeljnih bioetičkih načela autonomije, dobročinstva, neškodljivosti i pravednosti, može se zaključiti da razvoj i primjena umjetne inteligencije u psihoterapiji moraju biti usmjereni na očuvanje ljudskog dostojanstva i integriteta stručnog psihoterapijskog odnosa. To podrazumijeva jasno razgraničenje između tehnologije kao pomoćnog sredstva i psihoterapije kao interpersonalnog procesa koji se ne može u potpunosti automatizirati. Upravo u toj granici nalazi se ključ odgovorne i bioetički prihvatljive integracije umjetne inteligencije u psihoterapijsku praksu. Budući razvoj ovog područja ovisit će o sposobnosti zakonodavca, stručne zajednice i društva u cjelini da prepoznaju i očuvaju ono što psihoterapiju čini specifičnom, istodobno koristeći prednosti tehnološkog napretka bez ugrožavanja temeljnih vrijednosti na kojima počiva.

---

## Literatura

---

- Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., Househ, M. “Effectiveness and Safety of Using Chatbots to Improve Mental Health: Systematic Review and Meta-Analysis”, *Journal of Medical Internet Research*, 22, br. 7 (2019): 1-22. DOI: 10.2196/preprints.16021
- Akhil, P. J., Anithamol, B. “The Unseen Dilemma of AI in Mental Healthcare”, *AI & Society*, 40, br. 3 (2025): 1533-1535. DOI: 10.1007/s00146-024-01937-9
- Beauchamp, T. L., Childress, J. F. *Principles of Biomedical Ethics*, 8th ed. Oxford: Oxford University Press, 2019.
- Bordin, E. S. “The Generalizability of the Psychoanalytic Concept of the Working Alliance”, *Psychotherapy: Theory, Research & Practice*, 16, br. 3 (1979): 252-260.
- Ćurković, M. “Primjena chatbotova u širem okviru skrbi za mentalno zdravlje”, u: *Praktična primjena chatbotova u svakodnevnom (psihijatrijskom) radu* (ur. Arbanas, G.), Zagreb: Klinika za psihijatriju Vrapče, Sveučilište u Zagrebu Medicinski fakultet (2024): 34-46.
- D’Alfonso, S. “AI in Mental Health”, *Current Opinion in Psychology*, 36 (2020): 112-117. DOI: 10.1016/j.copsyc.2020.04.005
- Davenport, T., Kalakota, R. “The Potential for Artificial Intelligence in Healthcare”, *Future Healthcare Journal*, 6, br. 2 (2019): 94-98.
- Fitzpatrick, K. K., Darcy, A., Vierhile, M. “Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial”, *JMIR Mental Health*, 4, br. 2 (2017): 1-11. DOI: 10.2196/mental.7785
- Flückiger, C., Del Re, A. C., Wampold, B. E., Horvath, A. O. “The Alliance in Adult Psychotherapy: A Meta-Analytic Synthesis”, *Psychotherapy*, 55, br. 4 (2018): 316-340. DOI: 10.1037/pst0000172
- Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., Jeste, D. V. “Artificial Intelligence for Mental Health and Mental Illnesses: An Overview”, *Current Psychiatry Reports*, 21, br. 11 (2019): 1-26. DOI: 10.1007/s11920-019-1094-0
- Haque, M. R., Rubya, S. “For an App Supposed to Make Its Users Feel Better, It Sure Is a Joke – An Analysis of User Reviews of Mobile Mental Health Applications”, *Proceedings of the ACM on Human-*

*Computer Interaction*, 6: 1-29. New York: Association for Computing Machinery (2022). DOI: 10.1145/3555146

Horvath, A. O., Del Re, A. C., Flückiger, C., Symonds, D. “Alliance in Individual Psychotherapy”, *Psychotherapy*, 48, br. 1 (2011): 9-16. DOI: 10.1037/a0022186

Iftikhar, Z., Xiao, A., Ransom, S., Huang, J., Suresh, H. “How LLM Counselors Violate Ethical Standards in Mental Health Practice: A Practitioner-Informed Framework”, *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8, br. 2: 1311-1323. Madrid: Association for the Advancement of Artificial Intelligence (2025).

Inkster, B., Sarda, S., Subramanian, V. “An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study”, *JMIR mHealth and uHealth*, 6, br. 11 (2018): 1-14. DOI: 10.2196/12106

Iwaya, L. H., Babar, M. A., Rashid, A., Wijayarathna, C. “On the Privacy of Mental Health Apps: An Empirical Investigation and Its Implications for App Development”, *Empirical Software Engineering*, 28, br. 1 (2023): 1-42. DOI: 10.1007/s10664-022-10236-0

Koutsouleris, N., Hauser, T. U., Skvortsova, V., De Choudhury, M. “From Promise to Practice: Towards the Realisation of AI-Informed Mental Health Care”, *The Lancet Digital Health*, 4, br. 11 (2022): 829-840. DOI: 10.1016/S2589-7500(22)00153-4

Krajnović, A., Hordov, M., Pešić, Z. “Integrirana briga o zdravlju i holistički pristup – nova paradigma upravljanja zdravljem populacije uz primjenu nove tehnologije i umjetne inteligencije”, u: *10. međunarodna znanstvena konferencija CRODMA*, 183-204. Varaždin: Sveučilište u Zagrebu Fakultet organizacije i informatike (2025).

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., Coiera, E. “Conversational Agents in Healthcare: A Systematic Review”, *Journal of the American Medical Informatics Association*, 25, br. 9 (2018): 1248-1258. DOI: 10.1093/jamia/ocy072

Luxton, D. D. (2013). „Artificial intelligence in psychological practice: Current and future applications and implications”, *Professional Psychology: Research and Practice*, 45, br. 5 (2013): 1-9. DOI: 10.1037/a0034559

Malešević, A., Čartolovni, A. “Umjetna inteligencija i donošenje odluka u medicini: pogled ključnih dionika zdravstvenog sustava”, *Glasnik Hrvatskog katoličkog liječničkog društva*, 34, br. 4 (2024): 295-300.

- Rahsepar Meadi, M., Sillekens, T., Metselaar, S., van Balkom, A., Bernstein, J., Batelaan, N. “Exploring the Ethical Challenges of Conversational AI in Mental Health Care: Scoping Review”, *JMIR Mental Health*, 12, br. 1 (2025): 1-19. DOI: 10.2196/60432
- Minerva, F., Giubilini, A. “Is AI the Future of Mental Healthcare?”, *Topoi*, 42, br. 3 (2023): 809-817. DOI: 10.1007/s11245-023-09932-3
- Norcross, J. C., Lambert, M. J. “Psychotherapy Relationships That Work III”, *Psychotherapy*, 55, br. 4 (2018): 303-315. DOI: 10.1037/pst0000193
- Olawade, D. B., Wada, O. Z., Odetayo, A., David-Olawade, A. C., Asaolu, F., Eberhardt, J. “Enhancing Mental Health with Artificial Intelligence: Current Trends and Future Prospects”, *Journal of Medicine, Surgery, and Public Health*, 3 (2024): 1-10. DOI: 10.1016 /j.glmmedi.2024.100099
- Qiu, J., He, Y., Juan, X., Wang, Y., Liu, Y., Yao, Z., Wu, Y., Jiang, X., Yang, L., Wang, M. “EmoAgent: Assessing and Safeguarding Human-AI Interaction for Mental Health Safety”, u: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 11741-11756. Suzhou: ACL SIGDAT Special Interest Group on Linguistic Data & Corpus-based Approaches to Natural Language Processing (2025).
- Russell, S., Norvig, P. *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, New Jersey: Prentice-Hall, 1995.
- Shatte, A. B. R., Hutchinson, D. M., Teague, S. J. “Machine Learning in Mental Health: A Scoping Review of Methods and Applications”, *Psychol Med.*, 49, br. 9 (2019): 1426-1448. DOI: 10.1017/S0033291719000151
- Topol, E. J. “High-Performance Medicine: The Convergence of Human and Artificial Intelligence”, *Nature Medicine*, 25, br. 1 (2019): 44-56.
- Torous, J., Roberts, L. W. “Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust”, *JAMA Psychiatry*, 74, br. 5 (2017): 437-438.
- Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., Torous, J. B. “Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape”, *The Canadian Journal of Psychiatry*, 64, br. 7 (2019): 456-464. DOI: 10.1177/0706743719828977
- Vayena, E., Haeusermann, T., Adjekum, A., Blasimme, A. “Digital Health: Meeting the Ethical and Policy Challenges”, *Swiss Medical Weekly*, 148 (2018): 1-9. DOI: 10.4414/smw. 2018.14571.

*Pravni izvori*

Uredba (EU) 2024/1689 Europskog parlamenta i Vijeća od 13. lipnja 2024. o utvrđivanju usklađenih pravila o umjetnoj inteligenciji i o izmjeni uredaba (EZ) br. 300/2008, (EU) br. 167/2013, (EU) br. 168/2013, (EU) 2018/858, (EU) 2018/1139 i (EU) 2019/2144 te direktiva 2014/90/EU, (EU) 2016/797 i (EU) 2020/1828 (Akt o umjetnoj inteligenciji) (Tekst značajan za EGP) (Službeni list Europske unije L2024/1689)

Smjernice Komisije o definiciji sustava umjetne inteligencije utvrđenoj u Uredbi (EU) 2024/1689 (C(2025) 5053)

Smjernice Komisije o zabranjenim praksama u području umjetne inteligencije uspostavljenima Uredbom (EU) 2024/1689 (C(2025) 5052)

Uredba (EU) 2016/679 Europskog parlamenta i Vijeća od 27. travnja 2016. o zaštiti pojedinaca u vezi s obradom osobnih podataka i o slobodnom kretanju takvih podataka te o stavljanju izvan snage Direktive 95/46/EZ (Opća uredba o zaštiti podataka) (Tekst značajan za EGP) (Službeni list Europske unije L119/1)

Zakon o djelatnosti psihoterapije (Narodne novine 18/22)

Zakon o zaštiti osoba s duševnim smetnjama (Narodne novine 76/14)

Zakon o zaštiti prava pacijenata (Narodne novine 169/04, 37/08)